DESCRIPTION

Human Proteins Having Hydrophobic Domains and DNAs Encoding These Proteins

TECHNICAL FIELD

The present invention relates to human proteins having hydrophobic domains, DNAs encoding these proteins, and expression vectors for these DNAs as well as eukaryotic cells expressing these DNAs. The proteins of the present invention can be employed as pharmaceuticals or as antigens for preparing antibodies against these proteins. The human cDNAs of the present invention can be utilized as probes for genetic diagnosis and gene sources for gene therapy. Furthermore, the cDNAs can be utilized as gene sources for large-scale production of the proteins encoded by these cDNAs. Cells into which these genes are introduced to express secretory proteins or membrane proteins in large quantity can be utilized for detection of the corresponding receptors or ligands, screening of novel small molecule pharmaceuticals and the like.

BACKGROUND ART

Cells secrete many proteins extracellularly. These secretory proteins play important roles in the proliferation control, the differentiation induction, the material transport, the biophylaxis, and the like of the cells. Unlike intracellular proteins, the secretory proteins exert their actions outside the cells. Therefore, they can be administered in the intracorporeal manner such as the injection or the drip, so that they possess hidden

potentialities as pharmaceuticals. In fact, a number of human secretory proteins such as interferons, interleukins, erythropoietin, thrombolytic agents and the like have been currently employed as pharmaceuticals. In addition, secretory proteins other than those described above are undergoing clinical trials for developing their use as pharmaceuticals. It is believed that the human cells produce many unknown secretory proteins. Availability of these secretory proteins as well as genes encoding them is expected to lead to development of novel pharmaceuticals utilizing these proteins.

On the other hand, membrane proteins play important roles, as signal receptors, ion channels, transporters and the like in the material transport and the signal transduction through the cell membrane. Examples thereof include receptors for various cytokines, ion channels for the sodium ion, the potassium ion, the chloride ion and the like, transporters for saccharides and amino acids and the like. The genes for many of them have already been cloned. It has been clarified that abnormalities of these membrane proteins are involved in a number of previously cryptogenic diseases. Therefore, discovery of a new membrane protein is expected to lead to elucidation of the causes of many diseases, so that isolation of new genes encoding the membrane proteins has been desired.

Heretofore, due to difficulty in the purification from human cells, many of these secretory proteins and membrane proteins have been isolated by genetic approaches. A general method is the so-called expression cloning method, in which a cDNA library is introduced into eukaryotic cells to express cDNAs, and the cells secreting, or expressing on the surface of membrane, the protein having the activity of

interest are then screened. However, only genes for proteins with known functions can be cloned by using this method.

In general, a secretory protein or a membrane protein possesses at least one hydrophobic domain within the protein. After synthesis in the ribosome, such domain works as a secretory signal or remains in the phospholipid membrane to be entrapped in the membrane. Accordingly, if the existence of a highly hydrophobic domain is observed in the amino acid sequence of a protein encoded by a cDNA when the whole base sequence of the full-length cDNA is determined, it is considered that the cDNA encodes a secretory protein or a membrane protein.

OBJECTS OF THE INVENTION

The main object of the present invention is to provide novel human proteins having hydrophobic domains, DNAs encoding these proteins, and expression vectors for these DNAs as well as transformed eukaryotic cells that are capable of expressing these DNAs. This object as well as other objects and advantages of the present invention will become apparent to those skilled in the art from the following description with reference to the accompanying drawings.

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02539.

Fig. 2 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02770.

Fig. 3 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02869.

Fig. 4 illustrates the hydrophobicity/hydrophilicity

profile of the protein encoded by clone HP02956.

Fig. 5 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02962.

Fig. 6 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03014.

Fig. 7 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10608.

Fig. 8 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10609.

Fig. 9 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10611.

Fig. 10 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10617.

Fig. 11 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02837.

Fig. 12 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02991.

Fig. 13 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03063.

Fig. 14 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03091.

Fig. 15 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03092.

Fig. 16 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03116.

Fig. 17 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10618.

Fig. 18 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10619.

Fig. 19 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10622.

Fig. 20 illustrates the hydrophobicity/hydrophilicity

profile of the protein encoded by clone HP10625.

Fig. 21 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP02883.

Fig. 22 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03140.

Fig. 23 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10628.

Fig. 24 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10629.

Fig. 25 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10635.

Fig. 26 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10636.

Fig. 27 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10640.

Fig. 28 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10644.

Fig. 29 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10656.

Fig. 30 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10672.

Fig. 31 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03194.

Fig. 32 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03219.

Fig. 33 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03236.

Fig. 34 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03237.

Fig. 35 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03267.

Fig. 36 illustrates the hydrophobicity/hydrophilicity

profile of the protein encoded by clone HP03270.

Fig. 37 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03298.

Fig. 38 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10631.

Fig. 39 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10658.

Fig. 40 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10663.

Fig. 41 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03165.

Fig. 42 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03266.

Fig. 43 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP03287.

Fig. 44 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10665.

Fig. 45 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10669.

Fig. 46 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10670.

Fig. 47 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10671.

Fig. 48 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10673.

Fig. 49 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10675.

Fig. 50 illustrates the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10683.

SUMMARY OF THE INVENTION

As the result of intensive studies, the present inventors have successfully cloned cDNAs encoding proteins having hydrophobic domains from the human full-length cDNA bank, thereby completing the present invention. Thus, the present invention provides a human protein having hydrophobic domain(s), namely a protein comprising any one of an amino acid sequence selected from the group consisting of SEQ ID NOS: 1 to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130. Moreover, the present invention provides a DNA encoding the above-mentioned protein, exemplified by a cDNA comprising any one of a base sequence selected from the group consisting of SEQ ID NOS: 11 to 30, 41 to 60, 71 to 90, 101 to 120 and 131 to 150 as well as an expression vector that is capable of expressing such DNA by in vitro translation or in eukaryotic cells and a transformed eukaryotic cell that is capable of expressing such DNA and of producing the above-mentioned protein.

DETAILED DESCRIPTION OF THE INVENTION

The proteins of the present invention can be obtained, for example, by a method for isolating proteins from human organs, cell lines or the like, a method for preparing peptides by the chemical synthesis based on the amino acid sequence of the present invention, or a method for producing proteins by the recombinant DNA technology using the DNAs encoding the hydrophobic domains of the present invention. Among these, the method for producing proteins by the recombinant DNA technology is preferably employed. For example, the proteins can be expressed in vitro by preparing an RNA by in vitro transcription from a vector having the

cDNA of the present invention, and then carrying out in vitro translation using this RNA as a template. Alternatively, introduction of the translated region into a suitable expression vector by the method known in the art may lead to expression of a large amount of the encoded protein in prokaryotic cells such as *Escherichia coli*, *Bacillus subtilis*, etc., and eukaryotic cells such as yeasts, insect cells, mammalian cells, etc.

In the case where the protein of the present invention is produced by expressing the DNA by in vitro translation, the protein of the present invention can be produced in vitro by introducing the translated region of this cDNA into a vector having an RNA polymerase promoter, and then adding the vector to an in vitro translation system such as a rabbit reticulocyte lysate or a wheat germ extract, which contains an RNA polymerase corresponding to the promoter. The RNA polymerase promoters are exemplified by T7, T3, SP6 and the like. The vectors containing these RNA polymerase promoters are exemplified by pKA1, pCDM8, pT3/T7 18, pT7/3 19, pBluescript II and the like. Furthermore, the protein of the present invention can be expressed in the secreted form or the form incorporated in the microsome membrane when a canine pancreas microsome or the like is added to the reaction system.

In the case where the protein of the present invention is produced by expressing the DNA in a microorganism such as *Escherichia coli* etc., a recombinant expression vector in which the translated region of the cDNA of the present invention is introduced into an expression vector having an origin which is capable of replicating in the microorganism, a promoter, a ribosome-binding site, a cDNA-cloning site, a terminator and the like is constructed. After transformation

of the host cells with this expression vector, the resulting transformant is grown, whereby the protein encoded by the cDNA can be produced in large quantity in the microorganism. In this case, a protein fragment containing any translated region can be obtained by adding an initiation codon and a termination codon in front of and behind the selected translated region to express the protein. Alternatively, the protein can be expressed as a fusion protein with another protein. Only the portion of the protein encoded by the cDNA can be obtained by cleaving this fusion protein with a suitable protease. The expression vectors for *Escherichia coli* are exemplified by the pUC series, pBluescript II, the pET expression system, the pGEX expression system and the like.

In the case where the protein of the present invention is produced by expressing the DNA in eukaryotic cells, the protein of the present invention can be produced as a secretory protein, or as a membrane protein on the cell-membrane surface, by introducing the translated region of the cDNA into an expression vector for eukaryotic cells that has a promoter, a splicing region, a poly(A) addition site and the like, and then introducing the vector into the eukaryotic cells. The expression vectors are exemplified by pKA1, pED6dpc2, pCDM8, pSVK3, pMSG, pSVL, pBK-CMV, pBK-RSV, EBV vectors, pRS, pYES2 and the like. Examples of eukaryotic cells to be used in general include mammalian cultured cells such as monkey kidney COS7 cells, Chinese hamster ovary CHO cells and the like, budding yeasts, fission yeasts, silkworm cells, Xenopus oocytes and the like. Any eukaryotic cells may be used as long as they are capable of expressing the proteins of the present invention. The expression vector can be introduced into the eukaryotic cells by using a method

known in the art such as the electroporation method, the calcium phosphate method, the liposome method, the DEAE-dextran method and the like.

After the protein of the present invention is expressed in prokaryotic cells or eukaryotic cells, the protein of interest can be isolated from the culture and purified by a combination of separation procedures known in the art. Examples of the separation procedures include treatment with a denaturing agent such as urea or a detergent, sonication, enzymatic digestion, salting-out or solvent precipitation, dialysis, centrifugation, ultrafiltration, gel filtration, SDS-PAGE, isoelectric focusing, ion-exchange chromatography, hydrophobic chromatography, affinity chromatography, reverse phase chromatography and the like.

The proteins of the present invention also include peptide fragments (of 5 amino acid residues or more) containing any partial amino acid sequences in the amino acid sequences represented by SEQ ID NOS: 1 to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130. These peptide fragments can be utilized as antigens for preparation of antibodies. Among the proteins of the present invention, those having the signal sequences are secreted in the form of mature proteins after the signal sequences are removed. Therefore, these mature proteins shall come within the scope of the protein of the present invention. The N-terminal amino acid sequences of the mature proteins can be easily determined by using the method for the determination of cleavage site of a signal sequence [JP 8-187100 A]. Furthermore, some membrane proteins undergo the processing on the cell surface to be converted to the secreted forms. Such proteins or peptides in the secreted forms shall also come within the scope of the protein of the present invention. In the case where

sugar chain-binding sites are present in the amino acid sequences of the proteins, expression of the proteins in appropriate eukaryotic cells affords the proteins to which sugar chains are attached. Accordingly, such proteins or peptides to which sugar chains are attached shall also come within the scope of the protein of the present invention.

The DNAs of the present invention include all the DNAs encoding the above-mentioned proteins. These DNAs can be obtained by using a method for chemical synthesis, a method for cDNA cloning and the like.

The cDNAs of the present invention can be cloned, for example, from cDNA libraries derived from the human cells. The cDNAs are synthesized by using poly(A)$^+$ RNAs extracted from human cells as templates. The human cells may be cells delivered from the human body, for example, by the operation or may be the cultured cells. The cDNAs can be synthesized by using any method such as the Okayama-Berg method [Okayama, H. and Berg, P., Mol. Cell. Biol. 2: 161-170 (1982)], the Gubler-Hoffman method [Gubler, U. and Hoffman, J., Gene 25: 263-269 (1983)] and the like. However, it is desirable to use the capping method [Kato, S. et al., Gene 150: 243-250 (1994)], as exemplified in Examples, in order to obtain a full-length clone in an effective manner. In addition, commercially available human cDNA libraries can be utilized. The cDNAs of the present invention can be cloned from the cDNA libraries by synthesizing an oligonucleotide on the basis of base sequences of any portion in the cDNA of the present invention and screening the cDNA libraries using this oligonucleotide as a probe for colony or plaque hybridization according to a method known in the art. In addition, the cDNA fragments of the present invention can be prepared from an mRNA isolated from human cells by the RT-

PCR method in which oligonucleotides which hybridize with both termini of the cDNA fragment of interest are synthesized, which are then used as the primers.

The cDNAs of the present invention are characterized in that they comprise any one of the base sequences represented by SEQ ID NOS: 11 to 20, 41 to 50, 71 to 80, 101 to 110, and 131 to 140 or the base sequences represented by SEQ ID NOS: 21 to 30, 51 to 60, 81 to 90, 111 to 120, and 141 to 150. Table 1 summarizes the clone number (HP number), the cells from which the cDNA clone was obtained, the total base number of the cDNA, and the number of the amino acid residues of the encoded protein, for each of the cDNAs.

Table 1

SEQ ID NO	HP number	Cells	Base number	Number of amino acid residues
1, 11, 21	HP02539	Saos-2	4485	647
2, 12, 22	HP02770	HT-1080	1509	350
3, 13, 23	HP02869	KB	3059	206
4, 14, 24	HP02956	KB	2367	213
5, 15, 25	HP02962	KB	2355	595
6, 16, 26	HP03014	Liver	1024	264
7, 17, 27	HP10608	Saos-2	1237	343
8, 18, 28	HP10609	KB	1332	244
9, 19, 29	HP10611	KB	1932	303
10, 20, 30	HP10617	HT-1080	1124	160
31, 41, 51	HP02837	HT-1080	4473	1445
32, 42, 52	HP02991	KB	2630	582
33, 43, 53	HP03063	HT-1080	1472	410
34, 44, 54	HP03091	Liver	1652	483
35, 45, 55	HP03092	Liver	2112	607
36, 46, 56	HP03116	KB	1087	314
37, 47, 57	HP10618	HT-1080	1694	94
38, 48, 58	HP10619	HT-1080	1522	218
39, 49, 59	HP10622	Liver	1591	460
40, 50, 60	HP10625	Liver	1249	216
61, 71, 81	HP02883	KB	4027	392
62, 72, 82	HP03140	HT-1080	2495	497
63, 73, 83	HP10628	HT-1080	1617	417
64, 74, 84	HP10629	WERI-RB	3269	649
65, 75, 85	HP10635	WERI-RB	458	93
66, 76, 86	HP10636	HT-1080	1712	425
67, 77, 87	HP10640	WERI-RB	1055	149
68, 78, 88	HP10644	WERI-RB	1616	396
69, 79, 89	HP10656	PMA-U937	1860	350
70, 80, 90	HP10672	Thymus	783	153
91, 101, 111	HP03194	KB	3438	303

92, 102, 112	HP03219	PMA-U937	1144	283
93, 103, 113	HP03236	HT-1080	2339	488
94, 104, 114	HP03237	HT-1080	1765	182
95, 105, 115	HP03267	Liver	1418	184
96, 106, 116	HP03270	PMA-U937	1211	140
97, 107, 117	HP03298	PMA-U937	1099	153
98, 108, 118	HP10631	WERI-RB	3489	173
99, 109, 119	HP10658	HT-1080	931	75
100, 110, 120	HP10663	PMA-U937	1123	159
121, 131, 141	HP03165	KB	3234	636
122, 132, 142	HP03266	HT-1080	2490	318
123, 133, 143	HP03287	Thymus	1465	82
124, 134, 144	HP10665	HT-1080	917	247
125, 135, 145	HP10669	WERI-RB	1306	206
126, 136, 146	HP10670	WERI-RB	2022	432
127, 137, 147	HP10671	Thymus	1227	306
128, 138, 148	HP10673	Thymus	2210	555
129, 139, 149	HP10675	Thymus	1493	250
130, 140, 150	HP10683	PMA-U937	1264	174

The same clones as the cDNAs of the present invention can be easily obtained by screening the cDNA libraries constructed from the human cell lines or human tissues utilized in the present invention using an oligonucleotide probe synthesized on the basis of the base sequence of the cDNA provided in any one of SEQ ID NOS: 11 to 30, 41 to 60, 71 to 90, 101 to 120, and 131 to 150.

In general, the polymorphism due to the individual differences is frequently observed in human genes. Accordingly, any cDNA in which one or plural nucleotides are added, deleted and/or substituted with other nucleotides in SEQ ID NOS: 11 to 30, 41 to 60, 71 to 90, 101 to 120, and 131 to 150 shall come within the scope of the present

invention.

Similarly, any protein in which one or plural amino acids are added, deleted and/or substituted with other amino acids resulting from the above-mentioned changes shall come within the scope of the present invention, as long as the protein possesses the activity of the protein having any one of the amino acid sequences represented by SEQ ID NOS: 1 to 10, 31 to 40, 61 to 70, 91 to 100, and 121 to 130.

The cDNAs of the present invention also include cDNA fragments (of 10 bp or more) containing any partial base sequence in the base sequences represented by SEQ ID NOS: 11 to 20, 41 to 50, 71 to 80, 101 to 110, and 131 to 140 or in the base sequences represented by SEQ ID NOS: 21 to 30, 51 to 60, 81 to 90, 111 to 120, and 141 to 150. Also, DNA fragments consisting of a sense strand and an anti-sense strand shall come within this scope. These DNA fragments can be utilized as the probes for the genetic diagnosis.

In addition to the activities and uses described above, the polynucleotides and proteins of the present invention may exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use;

as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological

fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the

form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol.

149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6- Nordan, R. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark,S.C. and Turner, K.J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as

candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent

has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a

subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function in vivo on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating

autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the

transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected ex vivo with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection in vivo.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class

II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J.

Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J.J. and Brunswick, M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994;

Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to

stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complementary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and

Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is

not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and

in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head

trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon);

International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among

other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among

other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke)).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include,

without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987;

Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly

(such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Examples

The present invention is specifically illustrated in more detail by the following Examples, but Examples are not intended to restrict the present invention. The basic procedures with regard to the recombinant DNA and the enzymatic reactions were carried out according to the literature ["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989]. Unless otherwise stated, restriction enzymes and various modifying enzymes to be used were those available from Takara Shuzo. The buffer compositions and the reaction conditions for each of the enzyme reactions were as described in the attached instructions. The cDNA synthesis was carried out according to the literature [Kato, S. et al., Gene 150: 243-250 (1994)].

(1) Selection of cDNAs Encoding Proteins Having Hydrophobic Domains

The cDNA library of fibrosarcoma cell line HT-1080 (WO 98/11217), the cDNA library of osteosarcoma cell line Saos-2 (WO 97/33993), the cDNA library of epidermoid carcinoma cell line KB (WO 98/11217) and the cDNA library of liver tissue delivered by the operation (WO 98/21328) were used as the

cDNA libraries. Additionally, the cDNA libraries constructed from phorbol ester-stimulated histiocytic lymphoma cell line U937 (ATCC CRL 1593) mRNA, human retinoblastoma cell line WERI-RB (ATCC HTB 169) mRNA and human thymus mRNA (Clontech) were also used. Full-length cDNA clones were selected from the respective libraries and the whole base sequences thereof were determined to construct a homo-protein cDNA bank consisting of the full-length cDNA clones. The hydrophobicity/hydrophilicity profiles were determined for the proteins encoded by the full-length cDNA clones registered in the homo-protein cDNA bank by the Kyte-Doolittle method [Kyte, J. & Doolittle, R. F., J. Mol. Biol. 157: 105-132 (1982)] to examine the presence or absence of a hydrophobic region. A clone that has a hydrophobic region being assumed as a secretory signal or a transmembrane domain in the amino acid sequence of the encoded protein was selected as a clone candidate.

(2) Protein Synthesis by In Vitro Translation

The plasmid vector bearing the cDNA of the present invention was used for in vitro transcription/translation with a T_NT rabbit reticulocyte lysate kit (Promega). In this case, [^{35}S]methionine was added to label the expression product with a radioisotope. Each of the reactions was carried out according to the protocols attached to the kit. Two micrograms of the plasmid was subjected to the reaction at 30°C for 90 minutes in the reaction solution of a total volume of 25 μl containing 12.5 μl μ of T_NT rabbit reticulocyte lysate, 0.5 μl of a buffer solution (attached to the kit), 2 μl of an amino acid mixture (without methionine), 2 μl of [^{35}S]methionine (Amersham) (0.37 MBq/μl), 0.5 μl of T7 RNA polymerase, and 20 U of RNasin. The experiment in the presence of a membrane system was carried

out by adding 2.5 μl of a canine pancreas microsome fraction (Promega) to the reaction system. To 3 μl of the reaction solution was added 2 μl of the SDS sampling buffer (125 mM Tris-hydrochloride buffer, pH 6.8, 120 mM 2-mercaptoethanol, 2% SDS solution, 0.025% bromophenol blue and 20% glycerol) and the resulting mixture was heated at 95°C for 3 minutes and then subjected to SDS-polyacrylamide gel electrophoresis. The molecular weight of the translation product was determined by carrying out the autoradiography.

(3) Expression in COS7

Escherichia coli cells harboring the expression vector for the protein of the present invention were cultured at 37°C for 2 hours in 2 ml of the 2xYT culture medium containing 100 μg/ml of ampicillin, the helper phage M13KO7 (50 μl) was added, and the cells were then cultured at 37°C overnight. Single-stranded phage particles were obtained by polyethylene glycol precipitation from a supernatant separated by centrifugation. The particles were suspended in 100 μl of 1 mM Tris-0.1 mM EDTA, pH 8 (TE).

The cultured cells derived from monkey kidney, COS7, were cultured at 37°C in the presence of 5% CO_2 in the Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum. 1 x 10^5 COS7 cells were inoculated into a 6-well plate (Nunc, well diameter: 3 cm) and cultured at 37°C for 22 hours in the presence of 5% CO_2. After the medium was removed, the cell surface was washed with a phosphate buffer solution followed by DMEM containing 50 mM Tris-hydrochloride (pH 7.5) (TDMEM). A suspension containing 1 μl of the single-stranded phage suspension, 0.6 ml of the DMEM medium and 3 μl of TRANSFECTAM™ (IBF) was added to the cells and the cells were cultured at 37°C for 3 hours in the presence of 5% CO_2. After the sample solution was removed,

the cell surface was washed with TDMEM, 2 ml per well of DMEM containing 10% fetal calf serum was added, and the cells were cultured at 37°C for 2 days in the presence of 5% CO_2. After the medium was exchanged for a medium containing [^{35}S]cystine or [^{35}S]methionine, the cells were cultured for one hour. After the medium and the cells were separated each other by centrifugation, proteins in the medium fraction and the cell membrane fraction were subjected to SDS-PAGE.

(4) Clone Examples

<HP02539> (SEQ ID NOS: 1, 11, and 21)

Determination of the whole base sequence of the cDNA insert of clone HP02539 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 188-bp 5'-untranslated region, a 1944-bp ORF, and a 2353-bp 3'-untranslated region. The ORF encodes a protein consisting of 647 amino acid residues and there existed a putative secretory signal at the N-terminus and six putative transmembrane domains at the C-terminus. Figure 1 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse frizzled-1 (GenBank Accession No. AF054623). Table 2 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse frizzled-1 (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 90.4% in the entire

region.

Table 2

```
HP MAEEEAPKKSRAAGGGASWELCAGALSARLTEEGSGDAGGRRRPPVDPRRLARQLLLLLW
   ****.**..**** *  * ****.**..* .* * .*....***..**** *. ******
MM MAEEAAPSESRAA-GRLSLELCAEALPGRREEVGHEDTASHRRPRADPRRWASGLLLLLW
HP LLEAPLLLGVRAQAAGQGPGQGPGPGQQPPPPPQQQQSGQQYNGERGISVPDHGYCQPIS
   ***************    ** .*****.*****.**************.**********
MM LLEAPLLLGVRAQAA----GQVSGPGQQAPPPPQPQQSGQQYNGERGISIPDHGYCQPIS
HP IPLCTDIAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLCSMYAPVCTVL
   ******.*************************************************
MM IPLCTDMAYNQTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSAELKFFLCSMYAPVCTVL
HP EQALPPCRSLCERARQGCEALMNKFGFQWPDTLKCEKFPVHGAGELCVGQNTSDKGTPTP
   ************************************************************
MM EQALPPCRSLCERARQGCEALMNKFGFQWPDTLKCEKFPVHGAGELCVGQNTSDKGTPTP
HP SLLPEFWTSNPQHGGGGHRGGFPGGAGASERGKFSCPRALKVPSYLNYHFLGEKDCGAPC
   ********** ******.***.*****. ***********.********************
MM SLLPEFWTSNGQHGGGGYRGGYPGGAGTVERGKFSCPRALRVPSYLNYHFLGEKDCGAPC
HP EPTKVYGLMYFGPEELRFSRTWIGIWSVLCCASTLFTVLTYLVDMRRFSYPERPIIFLSG
   ********************************************.********** ***
MM EPTKVYGLMYFGPEELRFSRTWIGIWSVLCCASTLFTVLTYLVDMPRFSYPERPIISLSG
HP CYTAVAVAYIAGFLLEDRVVCNDKFAEDGARTVAQGTKKEGCTILFMMLYFFSMASSIWW
   ************************************.**********************
MM CYTAVAVAYIAGFLLEDRVVCNDKFAEDGARTVAQGTNKEGCTILFMMLYFFSMASSIWW
HP VILSLTWFLAAGMKWGHEAIEANSQYFHLAAWAVPAIKTITILALGQVDGDVLSGVCFVG
   ************************************************************.*
MM VILSLTWFLAAGMKWGHEAIEANSQYFHLAAWAVPAIKTITILALGQVDGDVLSGVCFLG
HP LNNVDALRGFVLAPLFVYLFIGTSFLLAGFVSLFRIRTIMKHDGTKTEKLEKLMVRIGVF
   ************************************************************
MM LNNVDALRGFVLAPLFVYLFIGTSFLLAGFVSLFRIRTIMKHDGTKTEKLEKLMVRIGVF
HP SVLYTVPATIVIACYFYEQAFRDQWERSWVAQSCKSYAIPCPHLQAGGGAPPHPPMSPDF
   ********************************************.***.**********
MM SVLYTVPATIVIACYFYEQAFRDQWERSWVAQSCKSYAIPCPHLQGGGGVPPHPPMSPDF
HP TVFMIKYLMTLIVGITSGFWIWSGKTLNSWRKFYTRLTNSKQGETTV
   *********                 **********************
MM TVFMIKYLM----------------TLNSWRKFYTRLTNSKQGETTV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA010020) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02770> (SEQ ID NOS: 2, 12, and 22)

Determination of the whole base sequence of the cDNA insert of clone HP02770 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 252-bp 5'-untranslated region, a 1053-bp ORF, and a 204-bp 3'-untranslated region. The ORF encodes a protein consisting of 350 amino acid residues and there existed two putative transmembrane domains. Figure 2 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 42 kDa that was somewhat larger than the molecular weight of 38,274 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human RING zinc finger protein (GenBank Accession No. AF037204). Table 3 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human RING zinc finger protein (ZN). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue

similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 56.0% in the entire region.

Table 3

```
HP          MHPAAFPLPVVVAAVLWGAAPTRGLIRATSDHNASMDFADLPALFGATLS
                                          .*** .*.**** **  *.
ZN MLLSIGMLMLSATQVYTILTVQLFAFLNLLPVEADILAYNFENASQTFDDLPARFGYRLP
HP QEGLQGFLVEAHPDNACSPIAPPPPAPVNGSVFIALLRRFDCNFDLKVLNAQKAGYGAAV
   .***.***....*.***.**.***    ....**.*.**.*****.******.*** **.
ZN AEGLKGFLINSKPENACEPIVPPPVKDNSSGTFIVLIRRLDCNFDIKVLNAQRAGYKAAI
HP VHNVNSNELLNMVWNSEEIQQQIWIPSVFIGERSSEYLRALFVYEKGARVLLVPDNTFPL
   ****.*..*..*  *. *. ..* ********.*.. *.. *.****....***. ..**
ZN VHNVDSDDLISMGSNDIEVLKKIDIPSVFIGESSANSLKDEFTYEKGGHLILVPEFSLPL
HP GYYLIPFTGIVGLLVLAMGAVMIARCIQHRKRLQRNRLTKEQLKQIPTHDYQKGDQYDVC
   .******  ***. .. .   **.. .*.*.* .**** *.***..*.*...***.****
ZN EYYLIPFLIIVGICLILIVIFMITKFVQDRHRARRNRLRKDQLKKLPVHKFKKGDEYDVC
HP AICLDEYEDGDKLRVLPCAHAYHSRCVDPWLTQTRKTCPICKQPVHRGPGDED-QEEETQ
   **************.***.****..*******.*.****.*** * ...**.* .....*
ZN AICLDEYEDGDKLRILPCSHAYHCKCVDPWLTKTKKTCPVCKQKVVPSQGDSDSDTDSSQ
HP GQEEGDEGEPRDHPASERTPLLGSSPTLPTSFGSLAPAPLVFPGPSTDPPLSPPSSPVIL
   ...* .* .*  .*
ZN EENEVTEHTPLLRPLASVSAQSFGALSESRSHQNMTESSDYEEDDNEDTDSSDAENEINE
HP V

ZN HDVVVQLQPNGERDYNIANTV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA434312) among ESTs. However, since they are partial sequences, it can not be judged whether or

not they encode the same protein as the protein of the present invention.

<HP02869> (SEQ ID NOS: 3, 13, and 23)

Determination of the whole base sequence of the cDNA insert of clone HP02869 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 229-bp 5'-untranslated region, a 621-bp ORF, and a 2209-bp 3'-untranslated region. The ORF encodes a protein consisting of 206 amino acid residues and there existed two putative transmembrane domains. Figure 3 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 22 kDa that was almost identical with the molecular weight of 22,367 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA278247) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02956> (SEQ ID NOS: 4, 14, and 24)

Determination of the whole base sequence of the cDNA insert of clone HP02956 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 68-bp 5'-untranslated region, a 642-bp ORF, and a 1657-bp 3'-untranslated region. The ORF encodes a protein consisting of 213 amino acid residues and there existed three putative transmembrane domains. Figure 4

depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 22 kDa that was almost identical with the molecular weight of 23,902 predicted from the ORF. When expressed in COS7 cells, an expression product of about 20 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human tetraspan NET-4 (GenBank Accession No. AF065389). Table 4 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human tetraspan NET-4 (TS). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 58.8% in the C-terminal region of 119 amino acid residues.

Table 4

```
HP                                                              MHY

TS MSGKHYKGPEVSCCIKYFIFGFNVIFWFLGITFLGIGLWAWNEKGVLSNISSITDLGGFD
HP YRYSNAKVSCWYKYLLFSYNIIFWLAGVVFLGVGLWAWSEKGVLSDLTKVTRMHGIDPVV

TS PVWLFLVVGGVMFILGFAGCIGALRENTFLLKFFSVFLGIIFFLELTAGVLAFVFKDWIK
HP LVLMVGVVMFTLGFAGCVGALRENICLLNFNQCCGAYGPEDWDLNVYFNCSGASYSREKC
                                  .. *****.*..**.**.****.... ***.*
TS DQLYFFINNNIRAYRDDIDLQNLIDFTQEYWQCCGAFGADDWNLNIYFNCTDSNASRERC
HP GVPFSCCVPDPAQKVVNTQCGYDVRIQLKSKWDESIFTKGCIQALESWLPRNIYIVAGVF
   *******. ***..*.*******.* . . . .  *.****....*.**. *. ****.*
TS GVPFSCCTKDPAEDVINTQCGYDARQKPEVDQQIVIYTKGCVPQFEKWLQDNLTIVAGIF
HP IAISLLQIFGIFLARTLISDIEAVKAGHHF
   *.*.******* **..*.******.*.
TS IGIALLQIFGICLAQNLVSDIEAVRASW
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T05279) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02962> (SEQ ID NOS: 5, 15, and 25)

Determination of the whole base sequence of the cDNA insert of clone HP02962 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 19-bp 5'-untranslated region, a 1788-bp ORF, and a 548-bp 3'-untranslated region. The ORF encodes a

protein consisting of 595 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 5 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 70 kDa that was somewhat larger than the molecular weight of 67,549 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 85 kDa to which sugar chains are presumably attached. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from alanine at position 23. In addition, there exist in the amino acid sequence of this protein four sites at which N-glycosylation may occur (Asn-Thr-Thr at position 75, Asn-Gln-Thr at position 153, Asn-Tyr-Thr at position 237 and Asn-Ser-Ser at position 360).

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human hypothetical protein KIAA0584 (GenBank Accession No. AB011156). Table 5 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human hypothetical protein KIAA0584 (KI). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 52.9% in the entire region.

Table 5

```
HP                MRAARAAPLLQLLLLLGPWLEAAGVAESPL--PAVVLAILARNAEHSL
                                      *..  .****  *.*..*.*****.*.*
KI LAWSLLLLSSALLREGCRARFVAERDSEDDGEEPVVFPESPLQSPTVLVAVLARNAAHTL
HP PHYLGALERLDYPRARMALWCATDHNVDNTTEMLQEWLAAVGDDYAAVVWRPEGEPRFYP
   **.** *******..***.* **********...*** .*   *  * *** .**  **
KI PHFLGCLERLDYPKSRMAIWAATDHNVDNTTEIFREWLKNVQRLYHYVEWRPMDEPESYP
HP DEEGPKHWTKERHQFLMELKQEALTFAR-NWGADYILFADTDNILTNNQTLRLLMGQGLP
   ** *****...* . .*.*.*.**  ** .* .***** *.**.*** ***.**.... .
KI DEIGPKHWPTSRFAHVMKLRQAALRTAREKW-SDYILFIDVDNFLTNPQTLNLLIAENKT
HP VVAPMLDSQTYYSNFWCGITPQGYYRRTAEYFPTKNRQRRGCFRVPMVHSTFLASLRAEG
   .*****.*.. **********.*.*.**..* ......* ***.********* .** *.
KI IVAPMLESRGLYSNFWCGITPKGFYKRTPDYVQIREWKRTGCFPVPMVHSTFLIDLRKEA
HP ADQLAFYPPHPNYTWPFDDIIVFAYACQAAGVSVHVCNEHRYGYMNVPVKSHQGLEDERV
   .*.*.*****..***.********.....**. ...** ..***. .*.*.**.*...
KI SDKLTFYPPHQDYTWTFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIPLKPHQTLQEDIE
HP NFIHLILEALVDGPRMQASAHVTRPSKRPSKIGFDEVFVISLARRPDRRERMLASLWEME
   *.**. .**..* *.*..*..*.  .* *.*.****.*.*.* ** ***.*** .*.* *
KI NLIHVQIEAMIDRPPMEPSQYVSVVPKYPDKMGFDEIFMINLKRRKDRRDRMLRTLYEQE
HP ISGRVVDAVDGWMLNSSAIRNLGVDLLPGYQDPYSGRTLTKGEVGCFLSHYSIWEEVVAR
   *. ..*.****  **.*....*....****.****.*.**.**.********.*.**..*
KI IEVKIVEAVDGKALNTSQLKALNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSVWKEVIDR
HP GLARVLVFEDDVRFESNFRGRLERLMEDVEAEKLSWDLIYLGRKQVN-PEKETAVEGLPG
   .*...**.******* .*. .* .**.......*.*.***.***...  * *.** ....
KI ELEKTLVIEDDVRFEHQFKKKLMKLMDNIDQAQLDWELIYIGRKRMQVKEPEKAVPNVAN
HP LVVAGYSYWTLAYALRLAGARKLLASQPLRRMLPVDEFLPIMFDQHPNEQYKAHFWPRDL
   ** *.******.*...*.**.**....*. .*********.*...** ..**... .***
KI LVEADYSYWTLGYVISLEGAQKLVGANPFGKMLPVDEFLPVMYNKHPVAEYKEYYESRDL
HP VAFSAQPLLAAPTHYAGDAEWLSDTETSSPWDDDSGRLISWSGSQ--KTLRSPRLDLTGS
    ****.***  ****.*....*******. **...   ..*. ..  *. . .*.  ...
KI KAFSAEPLLIYPTHYTGQPGYLSDTETSTIWDNETV-ATDWDRTHAWKSRKQSRIYSNAK
HP SGHSLQPQPRDEL
   ....*.*
KI NTEALPPPTSLDTVPSRDEL
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA358896) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03014> (SEQ ID NOS: 6, 16, and 26)

Determination of the whole base sequence of the cDNA insert of clone HP03014 obtained from cDNA library of human liver revealed the structure consisting of a 26-bp 5'-untranslated region, a 795-bp ORF, and a 203-bp 3'-untranslated region. The ORF encodes a protein consisting of 264 amino acid residues and there existed one putative transmembrane domain. Figure 6 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 31 kDa that was somewhat larger than the molecular weight of 28,471 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse WW domain-binding protein 1 (GenBank Accession No. U40825). Table 6 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse WW domain-binding protein 1 (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue

similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 85.1% in the entire region.

Table 6

```
HP                                                    MVASAKMGRAGTMAVAAELR
                                                                        **
MM MARASSRNSSEEAWGSLQAPQQQQSPAASSLEGAIWRRAGTQTRALDTILYHPQQSHLLR
HP ELCPGVNNQPYLCESGHCCGETGCCTYYYELWWFWLLWTVLILFSCCCAFRHRRAKLRLQ
   *******.******.**************************************************
MM ELCPGVNTQPYLCETGHCCGETGCCTYYYELWWFWLLWTVLILFSCCCAFRHRRAKLRLQ
HP QQQRQREINLLAYHGACHGAGPFPTGSLLDLRFLSTFKPPAYEDVVHRPGTPPPPYTVAP
   ********************** *********.**.***********.**********.*
MM QQQRQREINLLAYHGACHGAGPVPTGSLLDLRLLSAFKPPAYEDVVHHPGTPPPPYTVGP
HP GRPLTASSEQTCCSSSSSCPAHFEGTNVEGVSSHQSAPPHQEGEPGAGVTPASTPPSCRY
   * * *.*** * ***.***.**.**********.*** ******* **..*. .******
MM GYPWTTSSECTRCSSESSCSAHLEGTNVEGVSSQQSALPHQEGEPRAGLSPVHIPPSCRY
HP RRLTGDSGIELCPCPASGEGEPVKEVRVSATLPDLEDYSPCALPPESVPQIFPMGLSSSE
   ***************.*.****.**.*.**. *****.*******.**.*. ****.**
MM RRLTGDSGIELCPCPDSSEGEPLKEARASASQPDLEDHSPCALPPDSVSQVPPMGLASSC
HP GDIP
   *
MM GTSHK
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. W24575) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10608> (SEQ ID NOS: 7, 17, and 27)

Determination of the whole base sequence of the cDNA insert of clone HP10608 obtained from cDNA library of human osteosarcoma cell line Saos-2 revealed the structure consisting of a 23-bp 5'-untranslated region, a 1032-bp ORF, and a 182-bp 3'-untranslated region. The ORF encodes a protein consisting of 343 amino acid residues and there existed five putative transmembrane domains. Figure 7 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 37 kDa that was somewhat smaller than the molecular weight of 40,584 predicted from the ORF. When expressed in COS7 cells, an expression product of about 36 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T35406) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10609> (SEQ ID NOS: 8, 18, and 28)

Determination of the whole base sequence of the cDNA insert of clone HP10609 obtained from cDNA library of the human epidermoid carcinoma cell line KB revealed the structure consisting of a 38-bp 5'-untranslated region, a 735-bp ORF, and a 559-bp 3'-untranslated region. The ORF encodes a protein consisting of 244 amino acid residues and there existed one putative transmembrane domain at the N-terminus. Figure 8 depicts the hydrophobicity/hydrophilicity

profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 27 kDa that was almost identical with the molecular weight of 27,756 predicted from the ORF. When expressed in COS7 cells, an expression product of about 26 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Mycobacterium tuberculosis hypothetical protein Rv1147 (GenBank Accession No. Z95584). Table 7 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Mycobacterium tuberculosis hypothetical protein Rv1147 (MT). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 31.7% in the entire region.

Table 7

```
HP MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFS

MT                                MTSGAAASASRVDHPLFARIWPVVAAHEAEAIRAL
HP QIKGLTGASGKVALLELGCGTGANFQFYPPGC-RVTCLDPNPHFEKFLTKSMAENRHLQY
      *.* **.*  **.* *.*.**..** .  .*. ..*.*...   ..  *.. ...
MT RRENLAGLSGRV--LEVGAGVGTNFAYYPVAVEQVIAMEPEPRLAA-KARIAAADAPVPI
HP ERFVVAPGEDMRQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEP
     .  . *..*.  ....*.***.****** .*  **...*..**.** * . ****..
MT -VVTDKTVEEFRD--TETFDAVVCSLVLCSVSDPGAVLAHLRSLLRRGGELRYLEHVASA
HP YGSWAFMWQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKW--LPVGPH
    *... . . * .. * .....*   *.* ... .* *      *. .   *  ***.
MT -GARGRVQRFVDATFWPRLAGNCHTHRHTERAILDAGFVVDSSRREWAFPAWVPLPVSEL
HP IMGKAVK
    .*.* .
MT ALGRAHRT
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T60981) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10611> (SEQ ID NOS: 9, 19, and 29)

Determination of the whole base sequence of the cDNA insert of clone HP10611 obtained from cDNA library of the human epidermoid carcinoma cell line KB revealed the structure consisting of a 37-bp 5'-untranslated region, a 912-bp ORF, and a 983-bp 3'-untranslated region. The ORF

encodes a protein consisting of 303 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 9 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 31 kDa that was somewhat smaller than the molecular weight of 33,856 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 36 kDa. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from leucine at position 34. When expressed in COS7 cells, an expression product of about 35 kDa was observed in the membrane fraction.

The search of the protein data base using the amino acid sequence of the present protein revealed that the 218 amino acid residues at the C-terminus of the protein matched with the amino acid sequence of human glucosidase II (SWISS-PROT Accession No. Q06003). However, no similarity was observed at the N-terminal portion.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H14054) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10617> (SEQ ID NOS: 10, 20, and 30)

Determination of the whole base sequence of the cDNA insert of clone HP10617 obtained from cDNA library of the human fibrosarcoma cell line HT-1080 revealed the structure

consisting of a 72-bp 5'-untranslated region, a 483-bp ORF, and a 569-bp 3'-untranslated region. The ORF encodes a protein consisting of 160 amino acid residues and there existed four putative transmembrane domains. Figure 10 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight. When expressed in COS7 cells, an expression product of about 17 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H67672) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02837> (SEQ ID NOS: 31, 41, and 51)

Determination of the whole base sequence of the cDNA insert of clone HP02837 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 44-bp 5'-untranslated region, a 4338-bp ORF, and a 91-bp 3'-untranslated region. The ORF encodes a protein consisting of 1445 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 11 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 150 kDa that was almost identical with the molecular weight of 161,657 predicted from the ORF. Application of the (-3,-1) rule, a method for predicting the

cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from valine at position 22. In addition, there exist in the amino acid sequence of this protein 18 sites at which N-glycosylation may occur.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human α-2 macroglobulin (SWISS-PROT Accession No. P01023). Table 8 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human α-2 macroglobulin (MG). Therein, the marks of - and * represent a gap and an amino acid residue identical with that of the protein of the present invention, respectively. The both proteins shared a homology of 29.5% in the entire region.

Table 8

```
HP MQGPPLL--TAAHLLCVCTAALA-VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVT
   *    **      ** *     * *   *   *  *           *  **      **
MG MGKNKLLHPSLVLLLLVLLPTDASVSGKPQYMVLVP-SLLHTETTEKGCVLLSYLNETVT
HP VKAELLKTASN-LTVSVLEAE-GVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEIL
   * * *     *      ****  *     *     *    *         * * **
MG VSASLESVRGNRSLFTDLEAENDVLHCVAF---AVPKSSSNEEVMFLTVQVKGPTQ---E
HP FSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL--IKDPKS
   *   *          ** ****  *** * **** *     * *    *  *  * ***
MG FKKRTTVMVKNEDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHP-LNELIPLVYIQDPKG
HP NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQ-VNDQTYYQSFQVSEYVLPKFEVT
   * * ** * *   *     * *** *  *     **          * * * *******
MG NRIAQWQSFQLEGGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEHPFTVEEFVLPKFEVQ
HP LQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVT----LTFLPLSFWGKKKNITKTFKING
     *          *      ******* * **          *             *  *
MG VTVPKIITILEEEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSG
HP SANFSFNDEEMKNVMDSSNGLSEY-LDLSFPGPVEILTTVTESVTG----ISRNVSTNVF
    *                    **   *          ** *        * *      *
MG QLNSHGCFYQQVKTKVFQLKRKEYEMKLHTEAQIQEEGTVVELTGRQSSEITRTITKLSF
HP FK--QHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEY
    *   *      *     *          **    **         *     *
MG VKVDSHFRQGIPFFGQVRLVDGKGVPIPNKVIFIRGN-------EANYYSNATTDEHGLV
HP WSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK
      * * *                        *   *    **   * *   **
MG QFSINTTN-VMGTSLTVRVNYKDRSPCYGYQWVSEEHEEAHHTAYLVFSPSKSFVHLEPM
HP S--PSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVAVGKQ--NSTMF
   *        *          *    *       *       * *      **      *
MG SHELPCGHTQTVQAHYILNGGTLLGLKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHF
HP S-LTPENS-WTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVS
   *   *  *   * *    *     *  * * *  *  *     **  *  *      *
MG SISIPVKSDIAPVARLLIYAVLPTGDVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAH
HP LRISVT-QPDSIVGIVAVDKSVNLMNASNDITMENVVHEL-ELYNTG-------------
   **  **  * *     *** ** **          *   * *   **
MG LR--VTAAPQSVCALRAVDQSVLLMKPDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDC
```

```
HP ------YYLGMFMNSFAVFQE-CGLWVLTDANL---TKDYIDGVYDNAEYAERFMEENEG
          *  *           *       * *  *    *    *              *   **
MG INRHNVYINGITYTPVSSTNEKDMYSFLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEG
HP HIV----------DIHDFSLGSSPH---VRKHFPETWIWLDTNMGSRIYQEFEVTVPDSI
     *                       **    *** *******      *     *  ***** *
MG LRVGFYESDVMGRGHARLVHVEEPHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTI
HP TSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDA
   * * *  *  *** ***   *   * ******  *  ******** * *  *  ***
MG TEWKAGAFCLSEDAGLGISST-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKC
HP TEVKVIIEKSDKFDILMTSSE-----INATGHQ-QTLLVPSEDGATVLFPIRPTHL--GE
     * *  * *  *       *     * * * *     *       * *      *   *
MG IRVSVQLEASPAFLAVPVEKEQAPHCICANGRQTVSWAVTPKSLGNVNFTVSAEALESQE
HP IPITVTALSP--TASDAITQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPN
     *     *     *       **  ** **       **             **   ***
MG LCGTEVPSVPEHGRKDTVIKPLLVEPEGLEKETTFNSLL---CPSGGEVSEELSLKLPPN
HP TVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTD
    *  * *      ** **        *  ********** ****** ****    ***
MG VVEESARASVSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTP
HP NLKEKALSFMRQGYQRELLYQREDGSFSAFG--NYDPSGSTWLSAFVLRCFLEADPYIDI
     * **      **** * *   *** * **       * *** **** *  *  ** *
MG EVKSKAIGYLNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFI
HP DQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPN
   *         **   ** ** *   *         **    *** ***   **
MG DEAHITQALIWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTH
HP IDVQESIHFLES------EFSRGISDNYTLALITYALSSVG-SPKAKEALNMLTWRAEQE
     *      ***      *   * *  ** **  **    *   * ** * *   *
MG PVVRNALFCLESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKK
HP GGMQFW-----VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQ--TSE----GIPIMRW
       *                 *  *   *   * **         ***       *  *
MG DNSVHWERPQKPKAPVGHFYEPQAPSAEVEMTSYVLLAYLTAQPAPTSEDLTSATNIVKW
HP LSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSS-PSPVKFLIDTHN
     * *  *** ***** *** ***   *   * **     **  **     **  *  *
MG ITKQQNAQGGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNN
HP RLLLQTAELAVVQPTAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVA
   *****   *    *        * *    *    **                    *
```

```
MG RLLLQQVSL-PELPGEYSMKVTGEGCVYLQTSLKYN----ILPEKEEFPFALGVQTLPQT
HP VKENK-DDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVKKVEYDHG
     * *        * *    *    * **   *   ***      * *  *** *
MG CDEPKAHTSFQISLSVSYTGS-RSASNMAIVDVKMVSGF------IPLKPTVKMLE----
HP KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD
          *  *    *       ****        *    *    **           *
MG ----RSNHVSRTEVSSNHVLIYLDKVSNQTLSLFFTVLQDVP----VR----------D
HP LCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYFMELWL
   *           *               * * *
MG L---KPAIVKVYDYYETDEFAIAEYNAPCSKDL----GNA
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. W33075) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02991> (SEQ ID NOS: 32, 42, and 52)

Determination of the whole base sequence of the cDNA insert of clone HP02991 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 81-bp 5'-untranslated region, a 1749-bp ORF, and a 800-bp 3'-untranslated region. The ORF encodes a protein consisting of 582 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 12 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 66 kDa that was somewhat larger than the molecular weight of 64,244 predicted from the ORF. In

this case, the addition of a microsome led to the formation of a product of 78 kDa to which sugar chains are presumably attached. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from valine at position 27. In addition, there exist in the amino acid sequence of this protein seven sites at which N-glycosylation may occur (Asn-Gly-Thr at position 70, Asn-Gly-Thr at position 182, Asn-Gly-Ser at position 294, Asn-His-Thr at position 310, Asn-Gly-Thr at position 352, Asn-Glu-Thr at position 393 and Asn-Cys-Ser at position 407).

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse FKBP65-binding protein (GenBank Accession No. L07063). Table 9 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse FKBP65-binding protein (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 88.8% in the entire region.

Table 9

```
HP MFPAGPPSHSLLRLPLLQLLLLVVQAVGRGLGRASPAGGPLEDVVIERYHIPRACPREVQ
   ** .*..**.* *...*.**** .*...**********.**********************
MM MFLVGSSSHTLHRVRILPLLLL-LQTLERGLGRASPAGAPLEDVVIERYHIPRACPREVQ
HP MGDFVRYHYNGTFEDGKKFDSSYDRNTLVAIVVGVGRLITGMDRGLMGMCVNERRRLIVP
   *************************.************************************
MM MGDFVRYHYNGTFEDGKKFDSSYDRSTLVAIVVGVGRLITGMDRGLMGMCVNERRRLIVP
HP PHLGYGSIGLAGLIPPDATLYFDVVLLDVWNKEDTVQVSTLLRPPHCPRMVQDGDFVRYH
   *********.**********************.**** ..*****.*******..******
MM PHLGYGSIGVAGLIPPDATLYFDVVLLDVWNKADTVQSTILLRPPYCPRMVQNSDFVRYH
HP YNGTLLDGTSFDTSYSKGGTYDTYVGSGWLIKGMDQGLLGMCPGERRKIIIPPFLAYGEK
   *********.**.***.*******.**********************.**************
MM YNGTLLDGTGFDNSYSRGGTYDTYIGSGWLIKGMDQGLLGMCPGEKRKIIIPPFLAYGEK
HP GYGTVIPPQASLVFHVLLIDVHNPKDAVQLETLELPPGCVRRAGAGDFMRYHYNGSLMDG
   **************.***.*******.*********.****** ****************
MM GYGTVIPPQASLVFYVLLLDVHNPKDTVQLETLELPQGCVRRAVAGDFMRYHYNGSLMDG
HP TLFDSSYSRNHTYNTYIGQGYIIPGMDQGLQGACMGERRRITIPPHLAYGENGTGDKIPG
   ****************.*****************.*******.*****************
MM TLFDSSYSRNHTYNTYVGQGYIIPGMDQGLQGACIGERRRITVPPHLAYGENGTGDKIPG
HP SAVLIFNVHVIDFHNPADVVEIRTLSRPSETCNETTKLGDFVRYHYNCSLLDGTQLFTSH
   ******.**********.* ***.*****.*.****.*.***.************.**.**
MM SAVLIFDVHVIDFHNPSDPVEIKTLSRPPENCNETSKIGDFIRYHYNCSLLDGTRLFSSH
HP DYGAPQEATLGANKVIEGLDTGLQGMCVGERRQLIVPPHLAHGESGARGVPGSAVLLFEV
   **.**** ************ ************************.***************
MM DYEAPQEITLGANKVIEGLDRGLQGMCVGERRQLIVPPHLAHGENGARGVPGSAVLLFEV
HP ELVSREDGLPTGYLFVWHKDPPANLFEDMDLNKDGEVPPEEFSTFIKAQVSEGKGRLMPG
   *****************..**...*******************.******.*********
MM ELVSREDGLPTGYLFVWYQDPSTSLFEDMDLNKDGEVPPEEFSSFIKAQVNEGKGRLMPG
HP QDPEKTIGDMFQNQDRNQDGKITVDELKLKSDEDEERVHEEL
   ***.***.***************..*********.*******
MM QDPDKTISDMFQNQDRNQDGKITAEELKLKSDEDQERVHEEL
```

Furthermore, the search of the GenBank using the base

sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA308536) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03063> (SEQ ID NOS: 33, 43, and 53)

Determination of the whole base sequence of the cDNA insert of clone HP03063 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 88-bp 5'-untranslated region, a 1233-bp ORF, and a 151-bp 3'-untranslated region. The ORF encodes a protein consisting of 410 amino acid residues and there existed a putative transmembrane domain at the N-terminus. Figure 13 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 46 kDa that was almost identical with the molecular weight of 45,786 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the mouse AUP1 (GenBank Accession No. U41736). Table 10 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the mouse AUP1 (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 90.2% in the entire region.

Table 10

```
HP MELPSGPGPERLFDSHRLPGDCFLLLVLLLYAPVGFCLLVLRLFLGIHVFLVSCALPDSV
   ** *..**************.* ****.********.**********.**********
MM MEPPPAPGPERLFDSHRLPSDGFLLLALLLYAPVGLCLLVLRLFLGLHVFLVSCALPDSV
HP LRRFVVRTMCAVLGLVARQEDSGLRDHSVRVLISNHVTPFDHNIVNLLTTCSTPLLNSPP
   ****************************.**********************************
MM LRRFVVRTMCAVLGLVARQEDSGLRDHRVRVLISNHVTPFDHNIVNLLTTCSTPLLNSPP
HP SFVCWSRGFMEMNGRGELVESLKRFCASTRLPPTPLLLFPEEEATNGREGLLRFSSWPFS
   ************. * *******.**************************************
MM SFVCWSRGFMEMDRRVELVESLKKFCASTRLPPTPLLLFPEEEATNGREGLLRFSSWPFS
HP IQDVVQPLTLQVQRPLVSVTVSDASWVSELLWSLFVPFTVYQVRWLRPVHRQLGEANEEF
   ************************************************.*..*****..***
MM IQDVVQPLTLQVQRPLVSVTVSDASWVSELLWSLFVPFTVYQVRWLHPIRRQLGEESEEF
HP ALRVQQLVAKELGQTGTRLTPADKAEHMKRQRHPRLRPQSAQSSFPPSPGPSPDVQLATL
   **************.**************************.*****..*.**.****.**
MM ALRVQQLVAKELGQIGTRLTPADKAEHMKRQRHPRLRPQSVQSSFPSPPSPSSDVQLTTL
HP AQRVKEVLPHVPLGVIQRDLAKTGCVDLTITNLLEGAVAFMPEDITKGTQSLPTASASKF
   *.************.*******.************************.*.*.** *..**.**
MM AHRVKEVLPHVPLNVIQRDLARTGCVDLTITNLLEGAVAFMPEDVTEGSQSPPAPSAPKF
HP PSSGPVTPQPTALTFAKSSWARQESLQERKQALYEYARRRFTERRAQEAD
   **** .*********************************** **.****.
MM PSSGLATPQPTALTFAKSSWARQESLQERKQALYEYARRRFRERQAQEAE
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA131932) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03091> (SEQ ID NOS: 34, 44, and 54)

Determination of the whole base sequence of the cDNA insert of clone HP03091 obtained from cDNA library of human liver revealed the structure consisting of a 16-bp 5'-untranslated region, a 1452-bp ORF, and a 184-bp 3'-untranslated region. The ORF encodes a protein consisting of 483 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 14 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from leucine at position 34.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human OS-9 protein (SWISS-PROT Accession No. Q13438). Table 11 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human OS-9 protein (OS). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 27.8% in the N-terminal region of 281 amino acid residues. The positions of eight cysteines were conserved between the two proteins.

Table 11

```
HP MEEGGGGVRSLVPGGPVLLVLCGLLEASGGGRALPQLSDDIPFRVNWPGTEFSLPTTGVL
            .*...   **.* *** ... . ..  *. .    ...  . ..**. *
OS      MAAETLLSSLLGLLLL-GLLLPASLTGGVGSLNLEELSEMRYGIEILPLPVMGGQ
HP YKEDNYVIMTTAHKEKYKCILP----LVTSGDEEEEKDYKGPNPRELLEPLFKQSSCSYR
    .... **... .*..*.* **       .. ***. .*.**. .***.*. ....*  .
OS SQSSDVVIVSSKYKQRYECRLPAGAIHFQREREEETPAYQGPGIPELLSPM-RDAPCLLK
HP IESYWTYEVCHGKHIRQYHEEKETGQKINIHEYYLGNMLAKNLLFEKEREAEEKEKSNEI
   ....**** *.*.**.*** * ... * ..   ***        .... .      ..*.
OS TKDWWTYEFCYGRHIQQYHME-DSEIKGEV--LYLG-------YYQSAFD-----WDDET
HP PTKNIEGQMTPYYPVGMGNGTPCSLKQNRPRSSTVMYIC---HPESKHEILSVAEVTTCE
   .. . . ....*.. . ***. *.* ..***...*...*      * . * .*.*  .*.
OS AKASKQHRLKRYHSQTYGNGSKCDL-NGRPREAEVRFLCDEGAGISGDYIDRVDEPLSCS
HP YEVVILTPLLCSHPKYRFRASPV-NDIFCQ-SLPGSPFKPLTLRQLEQQEEILRVPFRRN
   * ..* ** **.**  * ..*.. ..*.*. **. . ... . ** . ..
OS YVLTIRTPRLCPHPLLRPPPSAAPQAILCHPSLQPEEYMAYVQRQADSKQYGDKIIEELQ
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA313678) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03092> (SEQ ID NOS: 35, 45, and 55)

Determination of the whole base sequence of the cDNA insert of clone HP03092 obtained from cDNA library of human liver revealed the structure consisting of a 19-bp 5'-untranslated region, a 1824-bp ORF, and a 269-bp 3'-untranslated region. The ORF encodes a protein consisting of

607 amino acid residues and there existed at least six putative transmembrane domains. Figure 15 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the rat liver-specific transport protein (GenBank Accession No. L27651). Table 12 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the rat liver-specific transport protein (RN). Therein, the marks of - and * represent a gap and an amino acid residue identical with that of the protein of the present invention, respectively. The both proteins shared a homology of 70.0% in the entire region.

Table 12

```
HP MGFEELLEQVGGFGPFQLRNVALLALPRVLLPLHFLLPIFLAAVPAHRCALPGAPANFSH
   **** **  ***********  * **** *** ***** * ****** ********* **
RN MGFEDLLDKVGGFGPFQLRNLVLMALPRMLLPMHFLLPVFMAAVPAHHCALPGAPANLSH
HP QDVWLEAHLPREPDGTLSSCLRFAYPQALPNTTLGEERQSRGELEDEPATVPCSQGWEYD
   ** ********* **  **********  ** *** *    ** * ** ***********
RN QDLWLEAHLPRETDGSFSSCLRFAYPQTVPNVTLGTEVSNSGEPEGEPLTVPCSQGWEYD
HP HSEFSSTIATESQVGIYIIHLEVECRWRQSPWEAAGRGLPWEEAEAAGLGRDKVSYSPSW
    *********
RN RSEFSSTIAT----------------------------------------------------
HP RESLGGLLSGMEWDLVCEQKGLNRAASTFFFAGVLVGAVAFGYLSDRFGRRRLLLVAYVS
              ****** * ***   ** ** *******  ********************
RN -----------EWDLVCQQRGLNKITSTCFFIGVLVGAVVYGYLSDRFGRRRLLLVAYVS
HP TLVLGLASAASVSYVMFAITRTLTGSALAGFTIIVMPLELEWLDVEHRTVAGVLSSTFWT
    ***** ****  * **  * *************** ***************** *  **
RN SLVLGLMSAASINYIMFVVTRTLTGSALAGFTIIVLPLELEWLDVEHRTVAGVISTVFWS
HP GGVMLLALVGYLIRDWRWLLLAVTLPCAPGILSLWWVPESARWLLTQGHVKEAHRYLLHC
   *** ********** ******* **** *** * ************** * **  *** *
RN GGVLLLALVGYLIRSWRWLLLAATLPCVPGIISIWWVPESARWLLTQGRVEEAKKYLLSC
HP ARLNGRPVCEDSFSQEAVSKVAAGERVVRRPSYLDLFRTPRLRHISLCCVVVWFGVNFSY
   * ****** * * ****   *   **   **********  ********  *********
RN AKLNGRPVGEGSLSQEALNNVVTMERALQRPSYLDLFRTSQLRHISLCCMMVWFGVNFSY
HP YGLSLDVSGLGLNVYQTQLLFGAVELPSKLLVYLSVRYAGRRLTQAGTLLGTALAFGTRL
   *** ************************  **  **  ***** ** *** ** *** *
RN YGLTLDVSGLGLNVYQTQLLFGAVELPSKIMVYFLVRRLGRRLTEAGMLLGAALTFGTSL
HP LVSSDMKSWSTVLAVMGKAFSEAAFTTAYLFTSELYPTVLRQTGMGLTALVGRLGGSLAP
   ***   *** * * * ***************************** ***** **** ***
RN LVSLETKSWITALVVVGKAFSEAAFTTAYLFTSELYPTVLRQTGLGLTALMGRLGASLAR
HP LAALLDGVWLSLPKLTYGGIALLAAGTALLLPETRQAQLPETIQDVERKSAPTSLQEEEM
   ********** ***  ****** ** ********  *************    * ***
RN LAALLDGVWLLLPKVAYGGIALVAACTALLLPETKKAQLPETIQDVERK----STQEE--
HP PMKQVQN

RN -----DV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AI016020) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03116> (SEQ ID NOS: 36, 46, and 56)

Determination of the whole base sequence of the cDNA insert of clone HP03116 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 32-bp 5'-untranslated region, a 945-bp ORF, and a 110-bp 3'-untranslated region. The ORF encodes a protein consisting of 314 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 16 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from lysine at position 20. In addition, there exist in the amino acid sequence of this protein three sites at which N-glycosylation may occur (Asn-Arg-Thr at position 167, Asn-Asn-Ser at position 200 and Asn-Ile-Ser at position 273).

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human Prostasin (SWISS-PROT Accession No. Q16651). Table 13 shows the comparison between amino acid sequences of the human protein of the present

invention (HP) and the human Prostasin (PR). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 39.8% in the entire region.

Table 13

```
HP      MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPW
          ..*.*  . .*   ** ....  *.  ..***  .  .**.**..*  *.***
PR MAQKGVLGPGQLGAVAILLYLGLLRSGTG-AEGAEAPCG-VAPQARITGGSSAVAGQWPW
HP QGSLRLWDSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAY
   * *.   . **** **.*..*.*.*****   ..  . ... *..*.   . *.    .
PR QVSITYEGVHVCGGSLVSEQWVLSAAHCF---PSEHHKEAYEVKLGA-HQLDSY---SED
HP YTRYFVSNIYLSPRYLGNSPY-DIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVT
    .   ...*   *.** ...  ****..** *.*....*.****.*.. .* *  .* **
PR AKVSTLKDIIPHPSYLQEGSQGDIALLQLSRPITFSRYIRPICLPAANASFPNGLHCTVT
HP GWGYIKEDEALPSPHTLQEVQVAIINNSMCNHLF-LKYSFRKDIF--GDMVCAGNAQGGK
   ***..  . .* .*..**...*..*... ** *. .. .  .  *  .****** ..***
PR GWGHVAPSVSLLTPKPLQQLEVPLISRETCNCLYNIDAKPEEPHFVQEDMVCAGYVEGGK
HP DACFGDSGGPLACNKNGLWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMS
   *** *******.*  .**** .*.**** .** .*******  * . .***. ...
PR DACQGDSGGPLSCPVEGLWYLTGIVSWGDACGARNRPGVYTLASSYASWIQSKVTELQPR
HP QPDPSWPLLFFPLLWALPLLGPV

PR VVPQTQESQPDSNLCGSHLAFSSAPAQGLLRPILFLPLGLALGLLSPWLSEH
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA159101) among ESTs. However, since they are partial sequences, it can not be judged whether or

not they encode the same protein as the protein of the present invention.

<HP10618> (SEQ ID NOS: 37, 47, and 57)

Determination of the whole base sequence of the cDNA insert of clone HP10618 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 215-bp 5'-untranslated region, a 285-bp ORF, and a 1194-bp 3'-untranslated region. The ORF encodes a protein consisting of 94 amino acid residues and there existed a putative transmembrane domain at the N-terminus. Figure 17 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 10 kDa that was almost identical with the molecular weight of 9,709 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA287125) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10619> (SEQ ID NOS: 38, 48, and 58)

Determination of the whole base sequence of the cDNA insert of clone HP10619 obtained from cDNA library of the human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 11-bp 5'-untranslated region, a 657-bp ORF, and a 854-bp 3'-untranslated region. The ORF encodes a protein consisting of 218 amino acid residues and there existed a putative transmembrane domain at the N-terminus.

Figure 18 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. Z43089) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10622> (SEQ ID NOS: 39, 49, and 59)

Determination of the whole base sequence of the cDNA insert of clone HP10622 obtained from cDNA library of the human liver revealed the structure consisting of a 43-bp 5'-untranslated region, a 1383-bp ORF, and a 165-bp 3'-untranslated region. The ORF encodes a protein consisting of 460 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 19 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from serine at position 17. In addition, there exist in the amino acid sequence of this protein four sites at which N-glycosylation may occur (Asn-Ser-Ser at position 23, Asn-Met-Ser at position 115, Asn-Glu-Thr at position 296 and Asn-Tyr-Thr at position 357).

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human angiopoietin-1 (GenBank

Accession No. U83508). Table 14 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human angiopoietin-1 (AN). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 28.2% in the entire region and a homology of 39.1% in the C-terminal region of 215 amino acid residues.

Table 14

```
HP                                                    MFTIKLLLFIVPLVISS

AN MTVFLSFAFLAAILTHIGCSNQRRSPENSGRRYNRIQHGQCAYTFILPEHDGNCRESTTD
HP RIDQDNSSFDSLSPEPKSRFAMLDDVKILANGLLQLGHGLKDF-VHKTKGQINDIFQKLN
                **. . ..*.... . ..  *  . *... *.. *.....* *. .
AN QYNTNALQRDAPHVEPDFSSQKLQHLEHVMENYTQWLQKLENYIVENMKSEMAQI-QQNA
HP IFDQSFYDLSLQTSEIKEEEKELRR-TTYKLQVKNEEVKNMSLELNSKLESLLEEKILLQ
   . ...   *.. ** ..... *. *. . ** *.. .    *...*..   ** ***
AN VQNHTATMLEIGTSLLSQTAEQTRKLTDVETQVLNQTSRLEIQLLENSLSTYKLEKQLLQ
HP QKVKYLE-EQLTNLIQNQPETPEHPEVTSLKTFVEKQDNSIKDLLQTVEDQYKQLNQQHS
   *. . *. .. ..*....  . *  . ..*.*. *...* ...*.   .   ..*..* .
AN QTNEILKIHEKNSLLEHKILEMEGKHKEELDTLKEEKEN-LQGLVTRQTYIIQELEKQLN
HP QIKEIENQLRRTSIQEPTEISLSSKPRAPRTTPFLQLNEIRNVKHDGIPAECTTIYNRGE
   . .. .. *..  . *  ..  .    .... .*  .. .. ... .. .*...*. *
AN RATTNNSVLQKQQL-ELMDTVHNLVNLCTKEGVLL--KGGKREEEKPFR-DCADVYQAGF
HP HTSGMYAIRPSN-SQVFHVYCDV-ISGSPWTLIQHRIDGSQNFNETWENYKYGFGRLDGE
   ..**.*.*  .* ..  .*.*.. ..*. **.**** *** .*. .*..** ***. .**
AN NKSGIYTIYINNMPEPKKVFCNMDVNGGGWTVIQHREDGSLDFQRGWKEYKMGFGNPSGE
HP FWLGLEKIYSIVKQSNYVLRIELEDWKDNKHYIEY-SFYLGNHETNYTLHLVAITGNVPN
   .*** * *..*..*..*.***** **..*. * .* .*..**.. ** *.* . **.. .
AN YWLGNEFIFAITSQRQYMLRIELMDWEGNRAYSQYDRFHIGNEKQNYRLYLKGHTGTAGK
HP AIP-ENKDLVFSTWDHKAKGHF-NCPEGYSGGWWWHDECGENNLNGKYNKPRAKSKPERR
   . .   ..  *** * .... . .*.   .****. *.** .****..    *  ....
AN QSSLILHGADFSTKDADNDNCMCKCALMLTGGWWF-DACGPSNLNGMFY--TAGQNHGKL
HP RGLSWKSQNGRLYSIKSTKMLIHPTDSESFE
   .*..*.  .*. **..**.*.*.* *
AN NGIKWHYFKGPSYSLRSTTMMIRPLDF
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for

example, Accession No. R86161) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10625> (SEQ ID NOS: 40, 50, and 60)

Determination of the whole base sequence of the cDNA insert of clone HP10625 obtained from cDNA library of the human liver revealed the structure consisting of a 133-bp 5'-untranslated region, a 651-bp ORF, and a 465-bp 3'-untranslated region. The ORF encodes a protein consisting of 216 amino acid residues and there existed two putative transmembrane domains. Figure 20 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R59052) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP02883> (SEQ ID NOS: 61, 71, and 81)

Determination of the whole base sequence of the cDNA insert of clone HP02883 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 191-bp 5'-untranslated region, a 1179-bp ORF, and a 2657-bp 3'-untranslated region. The ORF encodes a protein consisting of 392 amino acid residues and there existed three putative transmembrane domains. Figure 21 depicts the hydrophobicity/hydrophilicity profile, obtained

by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 40 kDa that was almost identical with the molecular weight of 43,381 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CET24F1.2 (GenBank Accession No. Z49912). Table 15 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Caenorhabditis elegans hypothetical protein CET24F1.2 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 35.4% in the N-terminal region of 178 amino acid residues.

Table 15

```
HP MEGVSALLARCPTAGLAGGLGVTACAAAGVLLYRIARRMKPTHTMVNCWFCNQDTLVPYG
                              *.  .**. * * . ..* *.****...* * *
CE                    MEVAAAVGVIASVPILYK-AIRPR-IKTSVECWFCRKSTKVEYQ
HP NRNCWDCPHCEQYNGFQENGDYNKPIPAQ-------YLEHLNHVVSSAPSLRDP-SQPQQ
   .**...** ******* *.****..**.*       * *. .   ..... *. .  .
CE QRNSFTCPSCEQYNGFTEDGDYNRRIPGQAWTTPKRYCEPGKMQSEKPSTFLDRFGGVNM
HP WVSSQVLLCKRCNHHQTTKIKQLAAFAPREEGRYDEEVEVYRHHLEQMYKLCRPCQAAVE
     ..   **. **  *.. ....*.*.* .*.*..**.* **..**.**.**..*  .*.
CE SPKASNGLCSECNLGQEIIMNKVAEFEPIDEDRWNEELEDYRYKLERMYQLCPRCTIQVH
HP YYIKHQNRQLRALLLSHQFKRREADQTHAQNFSSAVKSPVQVILLRALAFLACAFLLTTA
     .......  . **. ..* ..*
CE GKLEEDKKKY-SYLLKVKYKLKHAIGSTLREVMNNQKRSRRFFFAGGSTCEALHFGCLIS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. F11409) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03140> (SEQ ID NOS: 62, 72, and 82)

Determination of the whole base sequence of the cDNA insert of clone HP03140 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 29-bp 5'-untranslated region, a 1494-bp ORF, and a 972-bp 3'-untranslated region. The ORF encodes a protein consisting of 497 amino acid residues and there existed one putative transmembrane domain. Figure 22 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 51 kDa that was almost identical with the molecular weight of 54,245 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CELC50D2 (GenBank Accession No. AF040642). Table 16 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Caenorhabditis elegans hypothetical protein CELC50D2 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that

of the protein of the present invention, respectively. The both proteins shared a homology of 37.9% in the N-terminal region of 393 amino acid residues.

Table 16

```
HP MALWRGSAYAGFLALAVGCVFLLEPELPGSALRSLWSSLCLGPAPAPPGPVSPEGRLAAA
                                                    * *. .  .
CE          MFSETFVPSIFSYKHRLLHLSVLFFIVPYWYSYYNDQHRLSSYSVETAMFLS
HP WDALIVRPVRRWRRVAVGVNACVDVVLSGVKLLQALGLSPGNGKDHSILHSRNDLEEAFI
   *.  **.*  .......* *  **...***....**. . ..***...*.. .**...*
CE WERAIVKPGAMFKKAVIGFNCNVDLIVSGVRVVDALNTTCSEGKDQETLETLADLHQTFA
HP HFMWKGAAAERFFSDKETFHDIAQVASEFPGAQHYVGGNAALIGQKFAAN-SDLKVLLCG
   **. .******..*... *. ... ... . ..*..******.....*** .. .* * *
CE HFFQRGAAAERYMSSEDQFNLLVAESEASTRSHHHIGGNAALMADRIAANFPSTEVYLVG
HP PVGPRLHELLDDNVFVPPESLQEVDEFHLILEYQAGEEWGQLKAPHANRFIFSHDLSNGA
   *.*** ..**. .*  . ..    **.*.****..**  *.  ** ..*** ***  .*.
CE PIGPRSQALLHPSVKRTNSTRILKDELHVILEYKQGEILGDWVAPSSSRFITSHDHFSGS
HP MNMLEVFVSSLEEFQPDLVVLSGLHMMEGQSKELQRKRLLEVVTSISDIPTGIPVHLELA
   * ..*.* .....*.*****..*.*..* ****.....  .  .. .**. .*.****.
CE MVVMEMFFKAIAQFRPDLVVITGVHLLEFQSKEMRQEKMRLIKRNLLQIPPKVPIHLELG
HP SMTNRELMSSIVHQQVFPAVTSLGLNEQELLFLTQSASGPH-SSLSSWNGVPDVGMVSDI
   *... *..*. * ....* *.***.***** **.. *.*** .. .  .*. .* .* ..
CE SLAD-EIFSTDVINKILPYVDSLGINEQELTFLSHIANGPHMEEYPVQAGTVHVHKVVEM
HP LFWILKEHGR------SKSRASDLTRIHFHTLVYHILATVDGHWANQLAAVAAGARVAGT
   * *.**..**      ... .  *.***** *.***... ...*.*  *..*****.**
CE LHWLLKTYGRDPTGQIASKTGYRLSRIHFHCLTYHIMVSSGTDWSNLAAGLAAGARIAGR
HP QAC--ATETIDTSRVSLRAPQEFMTSHSEAGSRIVLNPNKPVVEWHREGISFHFTPVLVC
    .*  ...*.*.. ...*.*..*.
CE LSCNIGANTMDSELLEIRTPANFVLDKKIEKNYQFEAHKYMLTPFNIARCSTRLIRRKPP
HP KDPIRTVGLGDAISAEGLFYSEVHPHY
CE GGGILDEGVTFSDVHNVILNPTTRLPYPEEQLREHIEKTSSEIMKERNKIRYGTRKKKDS
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA356000) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10628> (SEQ ID NOS: 63, 73, and 83)

Determination of the whole base sequence of the cDNA insert of clone HP10628 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 66-bp 5'-untranslated region, a 1254-bp ORF, and a 297-bp 3'-untranslated region. The ORF encodes a protein consisting of 417 amino acid residues and there existed four putative transmembrane domains. Figure 23 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 46 kDa that was almost identical with the molecular weight of 45,461 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Schistosoma mansoni ATP-cassette family protein (GenBank Accession No. L26286). Table 17 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Schistosoma mansoni ATP-cassette family protein (SM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The

both proteins shared a homology of 39.5% in the C-terminal region of 294 amino acid residues.

Table 17

```
HP MLVHLFRVGIRGGPFPGRLLPPLRFQTFSAVRYSDGYRSSSLLRAVAHLRSQLWAHLPRA
SM                                MFSALCRRGFLTNKVSQFRSTYKCDHYNLKT
HP PLAPRWSPSAWCWVGGALLGPMVLSKHPHLCLVALCEAEEAPPASSTPHVVGSRFNWKLF
SM HIKPLKCSSSLRLTVGTGLFIALHSKISPESRIQTVQCEVDSYQTDQITFAKSGGIPRYI
HP WQFLHPHLLVLGVAVVLALGAALVNVQIPLLLGQLVEVVAKYTRDHVGSFMTESQNLSTH
     .. *. . *  *.. *. **..** *** **..*. ..... .* * *...    . .
SM GVLILPDCVYLFGAILGAFVAAVMNVYIPLYLGDFVSSLSRCVVTHEG-FVSAVYVPTLR
HP LLILYGVQGLLTFGYLVLLSHVGERMAVDMRRALFSSLLRYCQPQGAELGQDITFFDANK
   *   * .*.* ** *. **. ****** ** .**..*. *          **...**...
SM LCSSYLLQSLSTFLYIGLLGSVGERMARRMRIQLFRKLV-Y---------QDVAYFDVHS
HP TGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVAGCLVSLSMLSTRLTLLLMVATPALMG
   .*.**. ...***.******  ******. .**.*.. .*  .*. **  *.   * ..
SM SGKLVEIIGSDVQNFKSSFKQCISQGLRNGIQVVGSVFALLSISPTLTAALIGCLPCVFL
HP VGTLMGSGLRKLSCQCQEQIARAMGVADEALGNVRTVRAFAMEQREEERYGAELEACRCR
   .*.***..**..* . *.* .   .. ***....***...***.   .. . ...  .
SM IGSLMGTELRHISREVQSQNSLFASLIDEAFSHIRTVKSLAMEDFLINKINYNVDKAKML
HP AEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLTGGDLMSFLVASQTVQRL
   .*.*. **. ******...* .***.*..** *..  .*..*.*****...**.**
SM SEKLSFGIGSFQGLSNLTLNGVVLGVLYVGGHLMSRGELDAGHLMSFLATTQTLQRSLTQ
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. U66688) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10629> (SEQ ID NOS: 64, 74, and 84)

Determination of the whole base sequence of the cDNA insert of clone HP10629 obtained from cDNA library of human retinoblastoma cell line WERI-RB revealed the structure consisting of a 259-bp 5'-untranslated region, a 1950-bp ORF, and a 1060-bp 3'-untranslated region. The ORF encodes a protein consisting of 649 amino acid residues and there existed at least eight putative transmembrane domains. Figure 24 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CELF38B6 (GenBank Accession No. U40060). Table 18 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Caenorhabditis elegans hypothetical protein CELF38B6 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 39.1% in the C-terminal region of 445 amino acid residues.

Table 18

```
HP MIPNQHNAGAGSHQPAVFRMAVLDTDLDHILPSSVLPPFWAKLVVGSVAIVCFARSYDGD
CE                                          MKYAEINVNSGKHFRLNYKLHETS
HP FVFDDSEAIVNNKVAGVVGRADLLCALFFLLSFLGYCKAFRESNKEGAHSSTFWVLLSIF
CE TLGYHVVNIICHTVATLVFYKLGKQLEHIFDFFNIAFSASILFAVHPVHTEAVANITGRA
HP LGAVAMLCKEQGITVLGLNAVFDILVIGKFNVLEIVQKVLHKDKSLENLGMLRNGGLLFR
CE ELLMTIFSLAALILHVKNREINCKFVLLVILSTLSKEQGLMTIPIAICIDFLAHRSCRSN
HP MTLLTSGGAGMLYVRWRIMGTGPPAFTEVDNPASFADSMLVRAVNYNYYYSLNAWLLLCP
   ..* .. * ... **..***..* .*.. * .**.* .  .*.**. *
CE FVRMICLLVAIGFLRMMVNGFEAAKFTKLDNPTAFLNSKFYRMINYTYIWLYHAYLLVIP
HP WWLCFDWSMGCIPLIKSISDWRVIALAALWFCLIGLICQALCSEDGHKRRILTLGLGFLV
     ****.*****. *...  * . **... .  * .*   .     .. * .**.  . .
CE VNLCFDYSMGCISSITTM--WDLRALSPVLIFTIVIIGVKF----QNECRAFTLSSLMGI
HP IPFLPASNLFFRVGFVVAERVLYLPSIGYCVLLTFGFGALSKHTKKKKLIAAVVLGILFI
   *.******.** *** .********* *.*.* .. *  ** * *. .... ... .*.
CE ISFLPASNIFFTVGFSIAERVLYLPSAGFCLLCAIIFKKLSVHFKNADVLSITLILLLIS
HP NTLRCVLRSGEWRSEEQLFRSALSVCPLNAKVHYNIGKNLADKGNQTAAIRYYREAVRLN
   .* *   *****..*  *..*.***** ***.***.** *.*.*  ..* . *..*..*.
CE KTYR---RSGEWKTELSLYSSGLSVCPTNAKIHYNLGKVLGDNGLTKDAEKNYWNAIKLD
HP PKYVHAMNNLGNILKERNELQEAEELLSLAVQIQPDFAAAWMNLGIVQNSLKRFEAAEQS
   *.* .*.*****.*..... ..**.**. ** ..*.**.******* * .**.. .**.*
CE PSYEQALNNLGNLLEKSGDSKTAESLLARAVTLRPSFAVAWMNLGISQMNLKKYYEAEKS
HP YRTAIKHRRKYPDCYYNLGRLYADLNRHVDALNAWRNATVLKPEHSLAWNNMIILLDNTG
    ....  *.. ..* .*** **.  **.  *..**.*** ..*.** .*.*....**. .
CE LKNSLLIRPNSAHCLFNLGVLYQRTNRDEMAMSAWKNATRIDPSHSQSWTNLFVVLDHLS
HP NLAQAEAVGREALELIPNDHSLMFSLANVLGKSQKYKESEALFLKAIKANPNAASYHGNL
   . .*. ... .**. .**. .. . ...  .*  .....*. . .**. **... .*.**
CE QCSQVIDLSYQALSSVPNESRVHMQIGSCHAKHSNFTAAENHIKSAIDLNPTSVLFHANL
HP AVLYHRWGHLDLAKKHYEISLQLDPTASGTKENYGLLRRKLELMQKKAV
   ..**.* .. . *...* * *.**...  .*.*   *
CE GILYQRMSRHKEAESQYRIVLALDSKNIVAKQNLQKLEEHNCYNSTLP
```

Furthermore, the search of the GenBank using the base

sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA450191) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10635> (SEQ ID NOS: 65, 75, and 85)

Determination of the whole base sequence of the cDNA insert of clone HP10635 obtained from cDNA library of human retinoblastoma cell line WERI-RB revealed the structure consisting of a 65-bp 5'-untranslated region, a 282-bp ORF, and a 111-bp 3'-untranslated region. The ORF encodes a protein consisting of 93 amino acid residues and there existed two putative transmembrane domains. Figure 25 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 10 kDa that was almost identical with the molecular weight of 9,489 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA516481) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10636> (SEQ ID NOS: 66, 76, and 86)

Determination of the whole base sequence of the cDNA insert of clone HP10636 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure

consisting of a 179-bp 5'-untranslated region, a 1278-bp ORF, and a 255-bp 3'-untranslated region. The ORF encodes a protein consisting of 425 amino acid residues and there existed ten putative transmembrane domains. Figure 26 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. Z43270) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10640> (SEQ ID NOS: 67, 77, and 87)

Determination of the whole base sequence of the cDNA insert of clone HP10640 obtained from cDNA library of human retinoblastoma cell line WERI-RB revealed the structure consisting of a 52-bp 5'-untranslated region, a 450-bp ORF, and a 553-bp 3'-untranslated region. The ORF encodes a protein consisting of 149 amino acid residues and there existed at least two putative transmembrane domains. Figure 27 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 17 kDa that was almost identical with the molecular weight of 16,829 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Arabidopsis thaliana hypothetical

protein F27F23.14 (GenBank Accession No. AC003058). Table 19 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Arabidopsis thaliana hypothetical protein F27F23.14 (AT). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 46.5% in the entire region other than the N-terminal region.

Table 19

```
HP                                  METLYRVPFLVLECPNLKLKKPPWLHMPSAMTVYA
                                          ***  *  *.*.** *. . .**.*****
AT MAPRSDSQTGSSVSDGSDQSSMDPIFHLLRIVPFSFLRPPRLRLKIPS-FTLPSPMTVYA
HP LVVVSYFLITGGIIYDVIVEPPSVGSMTD-EHGHQRPVAFLAYRVNGQYIMEGLASSFLF
   *....***...*..********..**  * . *  ***.*.. *******.***.*.*.*
AT LILLTYFLVVSGFVYDVIVEPPGIGSTQDPTTGTIRPVVFMSGRVNGQYIIEGLSSGFMF
HP TMGGLGFIILDRSNAPNIPKLNRFLLLFIGFVCVLLSFFMARVFMRMKLPGYLMG
   ..**.*...** . . * .*  .      *  ..... *...*.*.*.****
AT VLGGIGIVMLDLALDKNKAKSVKASYAVAGVSSIVIAYVMSMLFIRIKIPGYLY
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. N34717) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10644> (SEQ ID NOS: 68, 78, and 88)

Determination of the whole base sequence of the cDNA insert of clone HP10644 obtained from cDNA library of the human retinoblastoma cell line WERI-RB revealed the structure consisting of a 221-bp 5'-untranslated region, a 1191-bp ORF, and a 204-bp 3'-untranslated region. The ORF encodes a protein consisting of 396 amino acid residues and there existed two putative transmembrane domains. Figure 28 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabiditis elegans hypothetical protein B0511.8 (GenBank Accession No. AF067608). Table 20 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the Caenorhabiditis elegans hypothetical protein B0511.8 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 31.3% in the region of 361 amino acid residues other than the N-terminal region and the C-terminal region.

Table 20

```
HS                                                  MAMIELGFGRQNFHPLKRKSSLLLKL
CE CDKNGQYLSVQEEIDAENKVQRKIAPGLNEKVLERVTQMLMKQEKSTETYMIWLKNLRVP
HS IAVVFAVLLFCEFLIYYLAIFQCNWPEVKTTASDGEQTTREPVLKAMFLADTHLLGEFLG
   * ... .... *..*...*.  *.**      .. .. . *. .**....******.. *
CE ILLAIILVVYNEYFIFFIAFSSCQWP------CKYGRCS-ESSVKAFMISDTHLLGKING
HS HWLDKLRREWQMERAFQTALWLLQPEVVFILGDIFDEGKWSTPEAWADDVERFQKMFRHP
   ******.*****  ..* .. *.  *.*.*.***..*****... .... .***.*.*
CE HWLDKLKREWQMYQSFWISTWIHSPDVTFFLGDLMDEGKWAGRPVFEAYAERFKKLF--G
HS SHVQLKVVAGNHDIGFHYEMNTYKVERFEKVFSSERLFSWKGINFVMVNSVALNGDGCGI
   .. .. ..*****.****..    *. *.. ....   . *.  ..* .....  ..
CE DNEKVITLAGNHDLGFHYAL----VQTFATHLTPT--VELKNYLLIMPETLEMFKKEFRR
HS CSETEAELIEVSHRLNCSREARG-SSR-CGPGPL--------LPTSAPVLLQHYPLYRRS
      .* .. . .  *  *.. .* ..* *  . *         *.. *..***.****.*
CE GLIDEMKIKKHRFVLINSMAMHGDGCRLCHEAELILEKIKSRNPKNRPIVLQHFPLYRKS
HS DANCSGEDAAPAEERDIPFKENYDVLSREASQKLLWWLQPRLVLSGHTHSAC-----EVH
   **.*.  *.... . .  ..*..*.**.*.* ...  *.*. *..****. *     ..
CE DAECDQVDEQHEIDLKEMYREQWDTLSKESSLQIIDSLNPKAVFGGHTHKMCKKKWNKTG
HS HGGRVPELSVPSFSWRNRNNPSFIMGSITPTDYTLSKCYLPREDVVLIIYC-GVVGFLVV
   ...   * .* ****** . *....  *. ..  .*.* **.* . ...*  * .*.*.
CE NSEYFYEYTVNSFSWRNGDVPAMLLVVIDGDNVLVSSCRLPSEILQIMVYIFGGIGILAK
HS LTLTHFGLLASPFLSGLNLLGKRKTR
      .
CE MYNDLITPAPLEWNVNNIAVCTAIILVMIINVVALIFTIFWCLRSKDEGGEIDSNGVVIN
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R88381) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10656> (SEQ ID NOS: 69, 79, and 89)

Determination of the whole base sequence of the cDNA insert of clone HP10656 obtained from cDNA library of the human lymphoma cell line U937 revealed the structure consisting of a 68-bp 5'-untranslated region, a 1053-bp ORF, and a 739-bp 3'-untranslated region. The ORF encodes a protein consisting of 350 amino acid residues and there existed two putative transmembrane domains. Figure 29 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 41 kDa that was almost identical with the molecular weight of 40,043 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 54 kDa to which sugar chains are presumably attached. In addition, there exist in the amino acid sequence of this protein four sites at which N-glycosylation may occur (Asn-Cys-Thr at position 148, Asn-Tyr-Thr at position 155, Asn-Gln-Thr at position 162 and Asn-Lys-Ser at position 190).

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA917816) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10672> (SEQ ID NOS: 70, 80, and 90)

Determination of the whole base sequence of the cDNA insert of clone HP10672 obtained from cDNA library of the

human thymus revealed the structure consisting of a 244-bp 5'-untranslated region, a 462-bp ORF, and a 77-bp 3'-untranslated region. The ORF encodes a protein consisting of 153 amino acid residues and there existed a putative secretory signal at the N-terminus and one putative transmembrane domain at the C-terminus. Figure 30 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. When expressed in COS cells, a product of 17 kDa was observed in the membrane fraction.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. N48700) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03194> (SEQ ID NOS: 91, 101, and 111)

Determination of the whole base sequence of the cDNA insert of clone HP03194 obtained from cDNA library of human epidermoid carcinoma cell line KB revealed the structure consisting of a 120-bp 5'-untranslated region, a 912-bp ORF, and a 2406-bp 3'-untranslated region. The ORF encodes a protein consisting of 303 amino acid residues and there existed four putative transmembrane domains. Figure 31 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the

protein was similar to the mouse hyperpolarization-activated cation channel HAC3 (GenBank Accession No. AJ225124). Table 21 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the mouse hyperpolarization-activated cation channel HAC3 (MM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 92.5% in the N-terminal region of 293 amino acid residues.

Table 21

```
HS MEAEQRPAAGASEGATPGLEAVPPVAPPPATAASGPIPKSGPEPKRRHLGTLLQPTVNKF
   **.*.******.*.***. *. **.**..* **** .*.*.******.************
MM MEEEARPAAGAGEAATPARET-PPAAPAQARAASGGVPESAPEPKRRQLGTLLQPTVNKF
HS SLRVFGSHKAVEIEQERVKSAGAWIIHPYSDFRFYWDLIMLLLMVGNLIVLPVGITFFKE
   ************************************************************
MM SLRVFGSHKAVEIEQERVKSAGAWIIHPYSDFRFYWDLIMLLLMVGNLIVLPVGITFFKE
HS ENSPPWIVFNVLSDTFFLLDLVLNFRTGIVVEEGAEILLAPRAIRTRYLRTWFLVDLISS
   ************************************************************
MM ENSPPWIVFNVLSDTFFLLDLVLNFRTGIVVEEGAEILLAPRAIRTRYLRTWFLVDLISS
HS IPVDYIFLVVELEPRLDAEVYKTARALRIVRFTKILSLLRLLRLSRLIRYIHQWEEIFHM
   ************************************************************
MM IPVDYIFLVVELEPRLDAEVYKTARALRIVRFTKILSLLRLLRLSRLIRYIHQWEEIFHM
HS TYDLASAVVRIFNLIGMMLLLCHWDGCLQFLVPMLQDFPPDCWVSINHMVVRSPHSSAFP
   ****************************************.*****.*.** .*
MM TYDLASAVVRIFNLIGMMLLLCHWDGCLQFLVPMLQDFPSDCWVSMNRMVNHSWGRQYSH
HS GPS

MM ALFKAMSHMLCIGYGQQAPVGMPDVWLTMLSMIVGATCYAMFIGHATALIQSLDSSRRQY
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AI571225) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03219> (SEQ ID NOS: 92, 102, and 112)

Determination of the whole base sequence of the cDNA insert of clone HP03219 obtained from cDNA library of human lymphoma cell line U937 revealed the structure consisting of a 55-bp 5'-untranslated region, a 852-bp ORF, and a 237-bp 3'-untranslated region. The ORF encodes a protein consisting of 283 amino acid residues and there existed four putative transmembrane domains. Figure 32 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human putative membrane protein 54TMp (GenBank Accession No. AF004876). Table 22 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the human putative membrane protein 54TMp (TM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 56.5% in the entire region.

Table 22

```
HS                      MADPHQLFDDTSSAQSRGYGAQRAPGGLSYPAASPT-PHAAF
                        .**..*****    * **..*  *** . ..*. . . . .
TM MAYHSGYGAHGSKHRARAAPDPPPLFDDT----SGGYSSQ--PGGYPATGADVAFSVNHL
HS LADPVSNMAMAYGSSLAAQGKELVDKNIDRFIPITKLKYYFAVDTMYVGRKLGLLFFPYL
   *.**..*.*******.*..**..*.*...**....****.***** **..***** ***
TM LGDPMANVAMAYGSSIASHGKDMVHKELHRFVSVSKLKYFFAVDTAYVAKKLGLLVFPYT
HS HQDWEVQYQQDTPVAPRFDVNAPDLYIPAMAFITYVLVAGLALGTQDRFSPDLLGLQASS
   **.***** .*.*..** *.********..*******.**.***.*.****..*** **.
TM HQNWEVQYSRDAPLPPRQDLNAPDLYIPTKAFITYVLLAGMALGIQKRFSPEVLGLCAST
HS ALAWLTLEVLAILLSLYLVTVNTDLTTIDLVAFLGYKYVGMIGGVLMGLLFGKIGYYLVL
   **.*...****.**.***.**..**.*..*.*. ******** .** *****. ***..*
TM ALVWVVMEVLALLLGLYLATVRSDLSTFHLLAYSGYKYVGMILSVLTGLLFGSDGYYVAL
HS GWCCVAIFVFMIRTLRLKILADAAAEGVPVRGARNQLRMYLTMAVAAAQPMLMYWLTFHL
   .* . *.. *..*.**   * .... * **  .*..*..***...** **...*******
TM AWTSSALMYFIVRSLRTAAL-GPDSMGGPV--PRQRLQLYLTLGAAAFQPLIIYWLTFHL
HS VR
   **
TM VR
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. H86659) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03236> (SEQ ID NOS: 93, 103, and 113)

Determination of the whole base sequence of the cDNA insert of clone HP03236 obtained from cDNA library of human

fibrosarcoma cell line HT-1080 revealed the structure consisting of a 252-bp 5'-untranslated region, a 1467-bp ORF, and a 620-bp 3'-untranslated region. The ORF encodes a protein consisting of 488 amino acid residues and there existed seven putative transmembrane domains. Figure 33 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein ZC513.5 (GenBank Accession No. U53155). Table 23 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the Caenorhabditis elegans hypothetical protein ZC513.5 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 39.5% in the intermediate region of 365 amino acid residues.

Table 23

```
HS MAGKGSSGRRPLLLGLLVAVATVHLVICPYTKVEESFNLQATHDLLYHWQDLEQYDHLEF
                                                        .*** .*
CE                                                    MKMKYDHSQF
HS PGVVPRTFLGPVVIAVFSSPAVYVLSLLEMSKFYSQLIVRGVLGLGVIFGLWTLQKEVRR
   ********.**. .*..***  ..... ...*..  *..*..***   ...  ....*.*
CE PGVVPRTFIGPISLAILSSPMSFIFRFWAIPKMWQLLLIRATLGLMNAMAFLYFARSVNR
HS HFGAMVATMFCWVTAMQFHLMFYCTRTLPNVLALPVVLLALAAWLRHEWARFIWLSAFAI
   .**  .*  .  .   *** .** .*.***..** .*.....  *   ... .. .. ..
CE KFGRETAMYLRLIMCTQFHYIFYMSRPLPNTFALILVMIVFERLLEGRYESAVRYATASV
HS IVFRVELCLFLGLLLL--LALGNRKV-SVVRALRHAVPAGILCLGLTVAVDSYFWRQLTW
   *.** ** *. * ..*  .  *. ** .   *.  .*. . .**.....***** .  *
CE ILFRCELVLLYGPIFLGYMISGRLKVFGFDGAIAIGVRIAAMCLAVSIPIDSYFWGRPLW
HS PEGKVLWYNTVLNKSSNWGTSPLLWYFYSALPRGLGCSLLFIPLG-LVDRRTHAPTVLAL
   ***.*...*.* *.* ..** *.********** *  . *..*** ***** .. .. ..
CE PEGEVMFFNVVENRSHEYGTQPFLWYFYSALPRCLLTTTLLVPLGLLVDRRLPQIVLPSV
HS GFMALYSLLPHKELRFIIYAFPMLNITAARGCSYLLNNYKKSWLYKAGSLLVIGHLVVNA
    *. ***.***********..*.. ..**  *. .* * .**.. .   . ** **..*.
CE IFIFLYSFLPHKELRFIIYVLPIFCLSAAVFCARMLINRHKSFFRMILFFGVILHLLANV
HS AYSATALYVSHFNYPGGVAMQ--RLHQLVPPQTDVLLHIDVAAAQTGVSRFLQVNSAWRY
    ...  * *.  **** *..  ....  ... * ..** * *****.***..*.**
CE LCTGMFLLVASKNYPGFDALNYLQFQNRFDAKKPVTVYIDNACAQTGVNRFLHINDAWT
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA744858) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03237> (SEQ ID NOS: 94, 104, and 114)

Determination of the whole base sequence of the cDNA insert of clone HP03237 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 101-bp 5'-untranslated region, a 549-bp ORF, and a 1106-bp 3'-untranslated region. The ORF encodes a protein consisting of 182 amino acid residues and there existed four putative transmembrane domains. Figure 34 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human intestinal membrane A4 protein (SWISS-PROT Accession No. Q04941). Table 24 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the human intestinal membrane A4 protein (IM). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 32.4% in the intermediate region of 111 amino acid residues.

Table 24

```
HS MWPPDPDPDPDPEPAGGSRPGPAVPGLRALLPARAFLCSLKGRLLLAESGLSFITFICYV
                                                ** **.**  *.....**.
IM                    MADSERLSAPGCWAACTNFSRTRKGILLFAEIILCLVILICF-
HS ASSASAFLTAPLLEFLLALYFLFADAMQLNDKWQGLCWPMMDFLRCVTAALIYFAISITA
   ..*.... . ...*..**  *. .   .*..* . . **  **.* ..**..*. .**..
IM SASTPGYSSLSVIEMILAAIFFVVYMCDLHTKIPFINWPWSDFFRTLIAAILYLITSIVV
HS IAKYSDGASKAAGVFGFFATIVFATDFYLIFNDVAKFLKQGDSADETTAHKTEEENSDSD
   ... .. .. .***.*..** .*. * *..*
IM LVERGNHSKIVAGVLGLIATCLFGYDAYVTFPVRQPRHTAAPTDPADGPV
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R14227) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03267> (SEQ ID NOS: 95, 105, and 115)

Determination of the whole base sequence of the cDNA insert of clone HP03267 obtained from cDNA library of human liver revealed the structure consisting of a 148-bp 5'-untranslated region, a 555-bp ORF, and a 715-bp 3'-untranslated region. The ORF encodes a protein consisting of 184 amino acid residues and there existed two putative transmembrane domains. Figure 35 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 21 kDa that was almost identical with the molecular

weight of 20,733 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human polyposis locus protein 1 (SWISS-PROT Accession No. Q00765). Table 25 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the human polyposis locus protein 1 (PL). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 59.1% in the entire region.

Table 25

```
HS MDGLRQRVEHFLEQRNLVTEVLGALEAKTGVEKRYLAAGAVTLLSLYLLFGYGASLLCNL
     .*.* ..**...* .*..*. *******.....* *...*..***.***********
PL    MRERFDRFLHEKNCMTDLLAKLEAKTGVNRSFIALGVIGLVALYLVFGYGASLLCNL
HS IGFVYPAYASIKAIESPSKDDDTVWLTYWVVYALFGLAEFFSDLLLSWFPFYYVGKCAFL
   *** **** ********.*.*** ********..*..*******..********. **.**
PL IGFGYPAYISIKAIESPNKEDDTQWLTYWVVYGVFSIAEFFSDIFLSWFPFYYMLKCGFL
HS LFCMAPRPWNGALMLYQRVVRPLFLRHHGAVDRIMNDLSGRALDAAAGITRNVKPSQTPQ
   *.****.* *** .**.*..**.**.*....*....**.... ..*..**...*
PL LWCMAPSPSNGAELLYKRIIRPFFLKHESQMDSVVKDLKDKSKETADAITKEAKKATVNL
HS PKDK

PL LGEEKKST
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for

example, Accession No. R09702) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03270> (SEQ ID NOS: 96, 106, and 116)

Determination of the whole base sequence of the cDNA insert of clone HP03270 obtained from cDNA library of human lymphoma cell line U937 revealed the structure consisting of a 132-bp 5'-untranslated region, a 423-bp ORF, and a 656-bp 3'-untranslated region. The ORF encodes a protein consisting of 140 amino acid residues and there existed four putative transmembrane domains. Figure 36 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 17 kDa that was somewhat larger than the molecular weight of 15,864 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Schizosaccharomyces pombe hypothetical protein (EMBL Accession No. AL031854). Table 26 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the Schizosaccharomyces pombe hypothetical protein (SP). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 43.4% in the entire region.

Table 26

```
HS      MSRFLNVLRSWLVMVSIIAMGNTLQSFRDHTFLYEKLYTGKPNLVNGLQARTFGI
            ... .* *.**. *. **.*** ... * ...* ...* *****.*****
SP MSQILAMLPDSLVAKWNVVVSVAALFNTVQSFLTPK-LTKRVY-SNTNEVNGLQGRTFGI
HS WTLLSSVIRCLCAIDIHNKTLYHITLWTFLLALGHFLSELFVYGTAAPTIGVLAPLMVAS
   *****...*  ** .* * ..* .   *. **  ***** ... *.. . *.*.*..*..
SP WTLLSAIVRFYCAYHITNPDVYFLCQCTYYLACFHFLSEWLLFRTTNLGPGLLSPIVVST
HS FSILGMLVGLRYLEVEPVSRQKKRN
    **
SP VSIWFMAKEKASILGIAA
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T30721) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03298> (SEQ ID NOS: 97, 107, and 117)

Determination of the whole base sequence of the cDNA insert of clone HP03298 obtained from cDNA library of human lymphoma cell line U937 revealed the structure consisting of a 182-bp 5'-untranslated region, a 462-bp ORF, and a 455-bp 3'-untranslated region. The ORF encodes a protein consisting of 153 amino acid residues and there existed at least one putative transmembrane domain. Figure 37 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 17.5 kDa that was almost identical with the molecular

weight of 17,360 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Schizosaccharomyces pombe hypothetical protein SPBC119.09c (EMBL Accession No. AL022117). Table 27 shows the comparison between amino acid sequences of the human protein of the present invention (HS) and the Schizosaccharomyces pombe hypothetical protein SPBC119.09c (SP). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 41.9% in the entire region other than the N-terminal region.

Table 27

```
HS                                MNVGVAHSEVNPNTRVMNSRGMWLTYALGVGLLHIVLL
                                             .* .* *... . .. *....
SP MGSSSSRRRSSSLVTKVPKPTIDDRLDQGSATNYNSNWVNYKGAWVIHIVLIAALRLIFH
HS SIPFFSVPVAWTLTNIIHNLGMYVFLHAVKGTPFETPDQGKARLLTHWEQLDYGVQFTSS
   .**  *  .******...  * ....* *.*****  . . .* ** ***** * *.*..
SP AIPSVSRELAWTLTNLTYMAGSFIMFHWVTGTPFEFNGGAYDR-LTMWEQLDEGNQYTPA
HS RKFFTISPIILYFLASFYTKYDPTHFILNTASLLSVLIPKMPQLHGVRIFGINKY
   **.. . ****..... **.*.   *..*. .*. *****.*..*  *****.*
SP RKYLLVLPIILFLMSTHYTHYNGWMFLVNIWALFMVLIPKLPAVHRKRIFGIQKLSLRDD
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA043039) among ESTs. However, since

they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10631> (SEQ ID NOS: 98, 108, and 118)

Determination of the whole base sequence of the cDNA insert of clone HP10631 obtained from cDNA library of the human retinoblastoma cell line WERI-RB revealed the structure consisting of a 226-bp 5'-untranslated region, a 522-bp ORF, and a 2741-bp 3'-untranslated region. The ORF encodes a protein consisting of 173 amino acid residues and there existed one putative transmembrane domain. Figure 38 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. W26443) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10658> (SEQ ID NOS: 99, 109, and 119)

Determination of the whole base sequence of the cDNA insert of clone HP10658 obtained from cDNA library of the human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 24-bp 5'-untranslated region, a 228-bp ORF, and a 679-bp 3'-untranslated region. The ORF encodes a protein consisting of 75 amino acid residues and there existed two putative transmembrane domains. Figure 39 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In

vitro translation resulted in formation of a translation product of 14 kDa or less that was almost identical with the molecular weight of 8,625 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. T85006) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10663> (SEQ ID NOS: 100, 110, and 120)

Determination of the whole base sequence of the cDNA insert of clone HP10663 obtained from cDNA library of the human lymphoma cell line U937 revealed the structure consisting of a 67-bp 5'-untranslated region, a 480-bp ORF, and a 576-bp 3'-untranslated region. The ORF encodes a protein consisting of 159 amino acid residues and there existed two putative transmembrane domains. Figure 40 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA336522) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03165> (SEQ ID NOS: 121, 131, and 141)

Determination of the whole base sequence of the cDNA insert of clone HP03165 obtained from cDNA library of human

epidermoid carcinoma cell line KB revealed the structure consisting of a 128-bp 5'-untranslated region, a 1911-bp ORF, and a 1195-bp 3'-untranslated region. The ORF encodes a protein consisting of 636 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 41 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 61 kDa that was smaller than the molecular weight of 72,033 predicted from the ORF. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from serine at position 33.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the human β-galactosidase (GenBank Protein ID No. AAA51822). Table 28 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the human β-galactosidase (GL). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 37.8% in the entire region.

Table 28

```
HP MTTWSLRRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDST
                  .*. .* .*..  *  .*    .  .* .... .  .*. .. .
GL                MPGFLVRILPLLLVLLLLGPTRGLRNATQRMFEIDYSRDSFLKDGQP
HP FWIFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLEAF
   *. ..***** **** **.*******  ***.. ******.*** .*...**.. *.* *
GL FRYISGSIHYSRVPRFYWKDRLLKMKMAGLNAIQTYVPWNFHEPWPGQYQFSEDHDVEYF
HP VLMAAEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMS
   . .* *.** **********.* ..****.***.. .. **..  .. .*** ... *..
GL LRLAHELGLLVILRPGPYICAEWEMGGLPAWLLEKESILLRSSDPDYLAAVDKWLGVLLP
HP RVVPLQYKRGGPIIAVQVENEYGSY-NKDPAYMPYVKKALEDR---GIVELLLTSDNKDG
   .. ** *..***.*.********** . * .*.....* . ..   ..* .  ....*.
GL KMKPLLLYQNGGPVITVQVENEYGSYFACDFDYLAFLQKRFRHHLGDDVVLFTTDGAHKTF
HP LSKGIVQGVLATINLQSTHELQLLTTFLF--NVQGTQPKMVMEYWTGWFDSWGGPHNILD
   *. * .**. .*... .. ..   ..**   . . . * .  *..***.* ** **....
GL LKCGALQGLYTTVDFGTGSNIT--DAFLSQRKCEPKGPLINSEFYTGWLDHWGQPHSTIK
HP SSEVLKTVSAIVDAGSSINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDY
   ...* ... .*.. *.*.***** *****.. ***     * .. ******* *.****
GL TEAVASSLYDILARGASVNLYMFIGGTNFAYWNGA--NSPYAAQPTSYDYDAPLSEAGDL
HP TAKYMKLRDFFGSISGIPLPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPIN
   *.**. **... ... .*  * *   **..*. .*     .. .**. *  *  . *  .
GL TEKYFALRNIIQKFEKVPEGPIPPSTPKFAYGKVTLEKLKTVGAALDILC-PSGPIK--S
HP MENLPVNGGNGQSFGYILYETSI----TSSGILSGH---VHDRGQVFVNTVSIGFLDYKT
   .  *.    . * .*..** *..    .... **..   ****. * *.... * *. ..
GL LYPLTFIQVK-QHYGFVLYRTTLPQDCSNPAPLSSPLNGVHDRAYVAVDGIPQGVLE-RN
HP TKIAVPLI-QGYTVLRILVENRGRVNYGENIDDQRKGLIGNLYLNDSPLKNFRIYSL---
   . *.. .. .. ..* .****.******. *.*  ***..** *... *... *..*
GL NVITLNITGKAGATLDLLVENMGRVNYGAYIND-FKGLVSNLTLSSNILTDWTIFPLDTE
HP DMKKSFFQRFG-------LDKWSSLPETPTLPAFFLGSLSI----SSTPCDTFLKLEGWE
   *  .* .  .*        . *.  ... *****..*..**    .. * ***... **.
GL DAVRSHLGGWGHRDSGHHDEAWAHNSSNYTLPAFYMGNFSIPSGIPDLPQDTFIQFPGWT
HP KGVVFINGQNLGRYW-NIGPQKTLYLPGP-WLSSGINQVIVFEETMAGPALQFTETPHLG
   ** *.*** ****** . *** **..* .  ..*. * ..*.*
GL KGQVWINGFNLGRYWPARGPQLTLFVPQHILMTSAPNTITVLELEWAPCSSDDPELCAVT
```

HP RNQYIK

GL FVDRPVIGSSVTYDHPSKPVEKRLMPPPPQKNKDSWLDHV

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA054017) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03266> (SEQ ID NOS: 122, 132, and 142)

Determination of the whole base sequence of the cDNA insert of clone HP03266 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 69-bp 5'-untranslated region, a 957-bp ORF, and a 1464-bp 3'-untranslated region. The ORF encodes a protein consisting of 318 amino acid residues and there existed one putative transmembrane domain at the N-terminus. Figure 42 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 34 kDa that was almost identical with the molecular weight of 35,363 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Arabidopsis thaliana putative ribotol dehydrogenase (GenBank Protein ID No. AAC23625). Table 29 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the

Arabidopsis thaliana putative ribotol dehydrogenase (AT). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 39.0% in the region of 483 residues other than the N-terminal region.

Table 29

```
HP MVELMFPLLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKE
                                      * ... .  ....**...*** *.*.
AT         MGIYGVMTGKKGKSGFGSASTAEDVTQAIDASHLTAIITGGTSGIGLEAARV
HP LAQRGARVYLACRDVEKGELVAKEIQTTTGNQQVLVRKLDLSDTKSIRAFAKGFLAEEKH
   ** ***.* .* *. . ..   . *   . *..*   ..*.*..**.*.*.. *** . .
AT LAMRGAHVIIAARNPKAANESKEMILQMNPNARVDYLQIDVSSIKSVRSFVDQFLALNVP
HP LHVLINNAGVMMCPYSKTADGFEMHIGVNHLGHFLLTHLLLEKLK-----ESAPSRIVNV
   *..********.**.. *.**.* ....**.******.***.*.*     .....****.
AT LNILINNAGVMFCPFKLTEDGIESQFATNHIGHFLLTNLLLDKMKSTARESGVQGRIVNL
HP SSLAH---HLGRIHFHNLQGEKFYNAGLAYCHSKLANILFTQELARRLKGSG--VTTYSV
   **.**   . . *.*.....   *..  ** .***.*.* ...*.***...*  .*. **
AT SSIAHTYTYSEGIKFQGINDPAGYSERRAYGQSKLSNLLHSNALSRRLQEEGVNITINSV
HP HPGTVQSELVRHSSFMRWMWWLFSF-FIKTPQQGAQTSLHCALTEGLEILSGNHFSDCHV
   *** * ..* *.*.* . ... ..* * *. .***.*. . **  .** ..*..*.**..
AT HPGLVTTNLFRYSGFSMKVFRAMTFLFWKNIPQGAATTCYVALHPDLEGVTGKYFGDCNI
HP AWVSAQARNETIARRLWDVSCDLLGLPID
   .  *  * *...* .*** *  *..
AT VAPSKFATNNSLADKLWDFSVFLIDSISK
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. D17020) among ESTs. However, since

they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP03287> (SEQ ID NOS: 123, 133, and 143)

Determination of the whole base sequence of the cDNA insert of clone HP03287 obtained from cDNA library of human thymus revealed the structure consisting of a 83-bp 5'-untranslated region, a 249-bp ORF, and a 1133-bp 3'-untranslated region. The ORF encodes a protein consisting of 82 amino acid residues and there existed one putative transmembrane domain at the N-terminus and one at the C-terminus, respectively. Figure 43 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of high molecular weight.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Schizosaccharomyces pombe hypothetical protein 9.0kDa (SWISS-PROT Accession No. O13825). Table 30 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Schizosaccharomyces pombe hypothetical protein 9.0kDa (SP). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 45.7% in the entire region.

Table 30

```
HP MAFTLYSLLQAALLCVNAIAVLHEERFLKNIGWGTDQGIGGFGE-EPGIKSQLMNLIRSV
     *.. ..* ..** .**.*.* *.*** .***. .... ***. . .***....***..
SP  MFGFGNILYVTLLLLNAVAILSEDRFLGRIGWSQSAAL-GFGDRQDTIKSRILHLIRAI
HP RTVMRVPLIIVNSIAIVLLLLFG
   ****  *** .*.*.** *..*
SP RTVMTFPLIAINTIVIVYNLVLG
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA853098) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10665> (SEQ ID NOS: 124, 134, and 144)

Determination of the whole base sequence of the cDNA insert of clone HP10665 obtained from cDNA library of human fibrosarcoma cell line HT-1080 revealed the structure consisting of a 31-bp 5'-untranslated region, a 744-bp ORF, and a 142-bp 3'-untranslated region. The ORF encodes a protein consisting of 247 amino acid residues and there existed a putative secretory signal at the N-terminus. Figure 44 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 28 kDa that was somewhat larger than the molecular weight of 25,320 predicted from the ORF. In this case, the addition of a microsome led to the formation

of a product of 27 kDa. Application of the (-3,-1) rule, a method for predicting the cleavage site of the secretory signal sequence, allows to expect that the mature protein starts from aspertic acid at position 26.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA055367) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10669> (SEQ ID NOS: 125, 135, and 145)

Determination of the whole base sequence of the cDNA insert of clone HP10669 obtained from cDNA library of human retinoblastoma cell line WERI-RB revealed the structure consisting of a 73-bp 5'-untranslated region, a 621-bp ORF, and a 612-bp 3'-untranslated region. The ORF encodes a protein consisting of 206 amino acid residues and there existed one putative transmembrane domain. Figure 45 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AF086533) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10670> (SEQ ID NOS: 126, 136, and 146)

Determination of the whole base sequence of the cDNA

insert of clone HP10670 obtained from cDNA library of human retinoblastoma cell line WERI-RB revealed the structure consisting of a 117-bp 5'-untranslated region, a 1299-bp ORF, and a 606-bp 3'-untranslated region. The ORF encodes a protein consisting of 432 amino acid residues and there existed seven putative transmembrane domains. Figure 46 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was similar to the Caenorhabditis elegans hypothetical protein CELM03F8.2 (GenBank Protein ID No. AAB65910). Table 31 shows the comparison between amino acid sequences of the human protein of the present invention (HP) and the Caenorhabditis elegans hypothetical protein CELM03F8.2 (CE). Therein, the marks of -, *, and . represent a gap, an amino acid residue identical with that of the protein of the present invention, and an amino acid residue similar to that of the protein of the present invention, respectively. The both proteins shared a homology of 39.6% in the N-terminal region of 376 residues.

Table 31

```
HP MDARWWAVVVLAAFPSLGAGGETPEAPPESWTQLWFFRFVVNAAGYASFMVPGYLLVQYF
                  .  .. ... **.  . .* .*. .   **..  .*. .*. *
CE            MDRSIMPIDSPARDKPPD--ELVWPLRLFLILLGYSTVATPAAILIYYV
HP RRKNYLETGRGLCFPLVKACVFGNEPKASDEVPLA---PRTEAAETTPMW-----QALKL
   **...  ... *...*.  . *. .  . . .* .   .*.*..... .       . *
CE RRNRHAFETPYLSIRLLLRS-FAVGNPEYQLIPTGEKQARKENDSIPQTRAQCINVIILL
HP LFCATGLQVSYLTWGVLQERVMTRSY-GATATSPGERFTDSQFLVLMNRVLALIVA--GL
   **  .*.**. .. ******..**.*  ... . ...*...***.. **..**...   *
CE LFFFSGIQVTLVAMGVLQERIITRGYRRSDQLEVEDKFGETQFLIFCNRIVALVLSLMIL
HP SCVLCKQPRHGAPMYRYSFASLSNVLSSWCQYEALKFVSFPTQVLAKASKVIPVMLMGKL
   .    ***.* .*.* .*..*.**..**********.******.. *****. .****.*
CE AKDWTKQPPHVPPLYVHSYTSFSNTISSWCQYEALKYVSFPTQTICKASKVVVTMLMGRL
HP VSRRSYEHWEYLTATLISIGVSMFLLSSGPEPRSSPAT--TLSGLILLAGYIAFDSFTSN
   *. ..*. .**  .  *..*.*.*****...  .*. *  ..**.**.***. **.** *
CE VRGQRYSWFEYGCGCTIAFGASLFLLSSSSKGAGSTITYTSFSGMILMAGYLLFDAFTLN
HP WQDALFAYK--MSSVQMMFGVNFFSCLFTVGSLLEQGALLEGTRFMGRHSEFAAHALLLS
   **.***. *  .*. ********** .. . **.***.* ....* . * .*. ...***
CE WQKALFDTKPKVSKYQMMFGVNFFSAILCAVSLIEQGTLWSSIKFGAEHVDFSRDVFLLS
HP ICSACGQLFIFYTIGQFGAAVFTIIMTLRQAFAILLSCLLYGHTVTVVGGLGVAVVFAAL
   ...* **.**. **..**. **..***.** *
CE LSGAIGQIFIYSTIERFGPIVFAVIMTIRQIFIRNTLIRAEDHRGVEMAPPPPPEPFRLK
HP LLRVYARGRLKQRGKKAVPVESPVQKV

CE FLSMIIAVIHI
```

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. Z46196) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the

present invention.

<HP10671> (SEQ ID NOS: 127, 137, and 147)

Determination of the whole base sequence of the cDNA insert of clone HP10671 obtained from cDNA library of human thymus revealed the structure consisting of a 74-bp 5'-untranslated region, a 921-bp ORF, and a 232-bp 3'-untranslated region. The ORF encodes a protein consisting of 306 amino acid residues and there existed a putative secretory signal at the N-terminus and one putative transmembrane domain at the intermediate region. Figure 47 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA357141) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10673> (SEQ ID NOS: 128, 138, and 148)

Determination of the whole base sequence of the cDNA insert of clone HP10673 obtained from cDNA library of the human thymus revealed the structure consisting of a 203-bp 5'-untranslated region, a 1668-bp ORF, and a 339-bp 3'-untranslated region. The ORF encodes a protein consisting of 555 amino acid residues and there existed one putative transmembrane domain. Figure 48 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product

of 65 kDa that was somewhat larger than the molecular weight of 61,781 predicted from the ORF.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. R96413) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10675> (SEQ ID NOS: 129, 139, and 149)

Determination of the whole base sequence of the cDNA insert of clone HP10675 obtained from cDNA library of the human thymus revealed the structure consisting of a 92-bp 5'-untranslated region, a 753-bp ORF, and a 648-bp 3'-untranslated region. The ORF encodes a protein consisting of 250 amino acid residues and there existed at least one putative transmembrane domain. Figure 49 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein.

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA356139) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

<HP10683> (SEQ ID NOS: 130, 140, and 150)

Determination of the whole base sequence of the cDNA insert of clone HP10683 obtained from cDNA library of the human lymphoma cell line U937 revealed the structure

consisting of a 25-bp 5'-untranslated region, a 525-bp ORF, and a 714-bp 3'-untranslated region. The ORF encodes a protein consisting of 174 amino acid residues and there existed one putative transmembrane domain. Figure 50 depicts the hydrophobicity/hydrophilicity profile, obtained by the Kyte-Doolittle method, of the present protein. In vitro translation resulted in formation of a translation product of 22 kDa that was somewhat larger than the molecular weight of 19,572 predicted from the ORF. In this case, the addition of a microsome led to the formation of a product of 24 kDa to which sugar chains are presumably attached. In addition, there exist in the amino acid sequence of this protein one site at which N-glycosylation may occur (Asn-Ile-Thr at position 27).

Furthermore, the search of the GenBank using the base sequences of the present cDNA has revealed the registration of sequences that shared a homology of 90% or more (for example, Accession No. AA482321) among ESTs. However, since they are partial sequences, it can not be judged whether or not they encode the same protein as the protein of the present invention.

INDUSTRIAL APPLICABILITY

The present invention provides human proteins having hydrophobic domains, DNAs encoding these proteins, and expression vectors for these DNAs as well as eukaryotic cells expressing these DNAs. Since all of the proteins of the present invention are secreted or exist in the cell membrane, they are considered to be proteins controlling the proliferation and/or the differentiation of the cells. Accordingly, the proteins of the present invention can be employed as pharmaceuticals such as carcinostatic agents

which act to control the proliferation and/or the differentiation of the cells, or as antigens for preparing antibodies against these proteins. The DNAs of the present invention can be utilized as probes for the genetic diagnosis and gene sources for the gene therapy. Furthermore, the DNAs can be utilized for large-scale expression of these proteins. Cells into which these genes are introduced to express these proteins, can be utilized for detection of the corresponding receptors or ligands, screening of novel small molecule pharmaceuticals and the like.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The

desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, Trends Pharmacol. Sci. 15(7): 250-254; Lavarosky et al., 1997, Biochem. Mol. Med. 62(1): 11-22; and Hampel, 1998, Prog. Nucleic Acid Res. Mol. Biol. 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, Bioessays 14(9): 629-633; Zwaal et al., 1993, Proc. Natl. Acad. Sci. USA 90(16): 7431-7435; Clark et al., 1994, Proc. Natl. Acad. Sci. USA 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour et al., 1988, Nature 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153;

5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s). Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75%

sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the Table 32 below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Table 32

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA : DNA	≥50	65℃; 1×SSC -or- 42℃; 1×SSC,50% formamide	65℃; 0.3×SSC
B	DNA : DNA	<50	T_B*; 1×SSC	T_B*; 1×SSC
C	DNA : RNA	≥50	67℃; 1×SSC -or- 45℃; 1×SSC,50% formamide	67℃; 0.3×SSC
D	DNA : RNA	<50	T_D*; 1×SSC	T_D*; 1×SSC
E	RNA : RNA	≥50	70℃; 1×SSC -or- 50℃; 1×SSC,50% formamide	70℃; 0.3×SSC
F	RNA : RNA	<50	T_F*; 1×SSC	T_F*; 1×SSC
G	DNA : DNA	≥50	65℃; 4×SSC -or- 42℃; 4×SSC,50% formamide	65℃; 1×SSC
H	DNA : DNA	<50	T_H*; 4×SSC	T_H*; 4×SSC
I	DNA : RNA	≥50	67℃; 4×SSC -or- 45℃; 4×SSC,50% formamide	67℃; 1×SSC
J	DNA : RNA	<50	T_J*; 4×SSC	T_J*; 4×SSC
K	RNA : RNA	≥50	70℃; 4×SSC -or- 50℃; 4×SSC,50% formamide	67℃; 1×SSC
L	RNA : RNA	<50	T_L*; 2×SSC	T_L*; 2×SSC
M	DNA : DNA	≥50	50℃; 4×SSC -or- 40℃; 6×SSC,50% formamide	50℃; 2×SSC
N	DNA : DNA	<50	T_N*; 6×SSC	T_N*; 6×SSC
O	DNA : RNA	≥50	55℃; 4×SSC -or- 42℃; 6×SSC,50% formamide	55℃; 2×SSC
P	DNA : RNA	<50	T_P*; 6×SSC	T_P*; 6×SSC
Q	RNA : RNA	≥50	60℃; 4×SSC -or- 45℃; 6×SSC,50% formamide	60℃; 2×SSC
R	RNA : RNA	<50	T_R*; 4×SSC	T_R*; 4×SSC

‡ : The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

† : SSPE (1×SSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH7.4) can be substituted for SSC (1×SSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R : The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10℃ less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(℃)=2(#of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(℃)=81.5 + 16.6($\log_{10}$[Na^+]) + 0.41 (%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1×SSC=0.165M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Current Protocols in Molecular Biology, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing

polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.